A Novel HARQ Design for RSMA Networks

Shaima Abidrabbu, Sawaira Rafaqat Ali, and Hüseyin Arslan, *Fellow, IEEE*

***Abstract*—Rate splitting multiple access (RSMA) has a variety of advantages, including higher throughput, enhanced capacity, and network efficiency, which makes it suitable for beyond fifth generation (5G) networks. Since the combining and splitting operations are performed on the transmitter side, one of the challenging aspects of RSMA is how a retransmission can be accomplished. In this article, an efficient hybrid automatic repeat request (HARQ) design for RSMA networks is presented in which three novel retransmission strategies are developed. More specifically, the main idea is choosing among RSMA, nonorthogonal multiple access (NOMA), and space division multiple access (SDMA) during the HARQ retransmission round, respectively, to do the retransmission. Therefore, a new link adaptation approach is proposed. The effectiveness of the proposed design is analyzed in terms of the average number of retransmissions, energy efficiency (EE), and achievable sum rate. Simulation results confirm the superiority of the proposed protocol in terms of packet error rate, EE, and sum-rate improvement up to 70%, 81%, and 76%, respectively, as compared to the conventional RSMA.**

***Index Terms*—Hybrid automatic repeat request (HARQ), latency, link adaptation, rate splitting multiple access (RSMA), retransmissions.**

I. Introduction

With the introduction of new dense applications and the Internet of Things (IoT) vision in fifth generation (5G), it is anticipated that the next 5G network will serve an unprecedented number of users along with achieving various key performance indicators, such as throughput, reliability, and latency. One of the promising ways to tackle this situation is switching from orthogonal to nonorthogonal design in multiple access techniques which allows simultaneous service of several users by the given available radio resources [1]. Multiple access systems can be classified into orthogonal and nonorthogonal based on the way of assigning the resources (frequency/time/code) [2]. Nonorthogonal approach dramatically increases spectrum efficiency by sharing the same radio resources with multiple users [3].

Recently, a general nonorthogonal multiple access (NOMA) framework, called rate splitting multiple access (RSMA) has been proposed [4]. RSMA is a powerful candidate for multiple access downlink multiantenna systems, and it includes both spatial division multiple access (SDMA) and NOMA as special cases [5]. RSMA relies on linearly precoded rate-splitting and successive interference cancellation (SIC) to decode part of the interference and treat the remaining part as noise [6]. In addition, RSMA has been shown to outperform both SDMA and NOMA ratewise in a wide range of network conditions, such as underloaded and overloaded regimes and user deployments with different conditions, such as diversity of channel directions, channel strengths, and qualities of channel state information (CSI) at the transmitter (CSIT) [5]. Due to these reasons, RSMA is one of the best candidates to provide efficient multiaccess for 5G and beyond networks. Furthermore, RSMA enables soft bridging of two extremes (NOMA and SDMA) by partially decoding interference and partially treating it as noise [6].

Hybrid automatic repeat request (HARQ) is one of the mechanisms that utilizes a feedback channel to do retransmission for recovery of signals that were not successfully received. In particular, the main idea behind HARQ is to combine the strengths of forward error correction (FEC) and automatic repeat request techniques to achieve reliable data transmission over a noisy channel [7]. Since HARQ improves throughput and reliability in the long-term evaluation (LTE) network, it has become a crucial component to investigate in the 3rd generation partnership project (3GPP) standard meetings [8]. In [7], the authors summarized how HARQ is applied to several wireless technologies, such as ultrareliable low-latency communication (URLLC), cooperative communications, multiple-input–multiple-output (MIMO), massive MIMO, NOMA, cognitive radio, caching, simultaneous wireless information, power transmission, and unmanned aerial vehicle (UAV) assisted communications.

A. Related Literature

HARQ schemes are widely used in wireless communication systems for reliable data transmission. It is employed in a variety of systems to reduce the bit error rate by increasing the transmission reliability, such as orthogonal frequency division multiple access (OFDMA) and NOMA. Regarding OFDMA, the resource allocation problem for HARQ-based OFDMA and single carrier FDMA networks is investigated in [9] with energy efficiency (EE) as the target parameter. For NOMA, the diversity order of three different HARQ-aided downlink NOMA systems is analyzed in [10]. Similarly, massive NOMA with HARQ is proposed in [11] for short packet communication under the same average received power constraints at a base station (BS). In terms of reliable and

Manuscript received 11 September 2023; revised 10 October 2023 and 24 October 2023; accepted 5 November 2023. Date of publication 13 November 2023; date of current version 26 March 2024. This work was supported by the Vestel Elektronik Sanayi ve Ticaret A.Ş. *(Corresponding author: Shaima Abidrabbu.)*

Shaima Abidrabbu is with the School of Engineering and Natural Science, Istanbul Medipol University, 34810 Istanbul, Turkey, and also with the IPR and License Agreements Department, Vestel Electronics, 45030 Manisa, Turkey (e-mail: shaima.abidrabbu@std.medipol.edu.tr).

Sawaira Rafaqat Ali and Hüseyin Arslan are with the School of Engineering and Natural Science, Istanbul Medipol University, 34810 Istanbul, Turkey (e-mail: sawaira.ali@std.medipol.edu.tr; huseyinarslan@medipol.edu.tr).

Digital Object Identifier 10.1109/JIOT.2023.3332096

secure transmission probability, HARQ enhances the secrecy performance in maximum ratio combining (MRC) and selection combing schemes, particularly in low-transmit power regions as presented in [12]. Additionally, in [13], the HARQ-based cooperative NOMA network is studied, where better outage performance is achieved by HARQ with storage than without storage case. Moreover, nonorthogonal HARQ (N-HARQ) outperforms the conventional HARQ in both additive white Gaussian noise (AWGN) and Rayleigh fading channels in terms of packet delivery latency in [14].

The aforementioned literature shows the capability of HARQ in terms of enhancing reliability based on different multiaccess networks. Reliability is a concern in RSMA networks and there are limited works in this direction. More specifically, in [15], the authors demonstrate that the uplink RSMA achieves more reliable communication with finite block length codes compared to uplink NOMA. In [16], RSMA in a short packet communication is investigated and it is proven that RSMA achieves the same maximum-minimum fairness rate as NOMA and SDMA with smaller block lengths. However, HARQ-based RSMA has not been studied yet due to the operations of splitting and combining on the transmitter side. As a result, a low-complexity and compatible HARQ design is a crucial step to enhance the reliability of RSMA networks. Motivated by this, we take a step toward developing a fully innovative HARQ strategy for RSMA networks. Moreover, the suggested approach also naturally offers a novel link adaptation based on various multiaccessing schemes by taking into consideration the RSMA characteristic of transitioning between NOMA and SDMA.

B. Problem Statement

Recently, the HARQ protocol has received great interest among researchers due to its direct relation with throughput, reliability, and latency. On the other hand, RSMA is one of the candidates for beyond 5G multiple access techniques. The process of retransmission-based RSMA networks is not as forward as the conventional HARQ schemes because the messages of various users are split and combined before being transmitted. Notably, when a user requests a retransmission in RSMA, it remains uncertain which data should be delivered again (common, private, or the entire packet). Hence, HARQ cannot be directly applied to RSMA due to its unique structure [17]. Therefore, there is a need to propose a HARQ protocol for RSMA networks.

In this article, an efficient design for HARQ-based RSMA networks is proposed in which three retransmission approaches are developed based on choosing among RSMA, NOMA, and SDMA during a HARQ retransmission round. To the best of the author's knowledge, this is the first work to design a HARQ protocol for RSMA networks. This study presents a qualitative leap not only in terms of proposing a novel HARQ protocol but also in providing RSMA to be a part of standards. Based on the simulation results, the suggested protocol significantly improved the achievable sum-rate and decreases packet error rate (PER) of the network when compared to the same network without HARQ and with existing HARQ approaches.

C. Contributions

In this article, we take a step toward merging the benefits of HARQ with multiuser down-link RSMA and shed light on its relation with average number of retransmissions and reliability performance. The main contributions of this article are as follows.

1) Our work is the first to provide the integration of HARQ protocol with RSMA networks. This article presents an efficient HARQ design for RSMA networks, as well as three unique retransmission strategies.
2) A new link adaptation is proposed which correspondingly chooses from the transmission topologies; NOMA, SDMA, and RSMA during a HARQ round because of the inherent flexibility that RSMA offers to a system that is bridging SDMA and NOMA. This is different compared to the conventional modulation and coding approach-based link adaptation.
3) To provide further insights, the proposed approach is evaluated in terms of several performance metrics, namely, PER, the average number of retransmissions and achievable sum rate improvement. For the proposed HARQ-RSMA approach, we develop a closed-form for the average number of retransmissions, and EE.
4) Based on the simulation results, the suggested protocol can provide a higher data rate and greater EE, compared to the conventional RSMA network without HARQ taking into account the AWGN channel, Rayleigh fading channel, average number of retransmission, and the size of data.

D. Organization

The remainder of this article is structured as follows. The system model is discussed in Section II, while Section III introduces the proposed retransmission protocol. In Section IV, the suggested strategy is examined. Section V provides the analytical and simulation results. Section VI presents the conclusion as well as recommendations for further investigation.

Notations: Bold uppercase $\mathbf{A}$, bold lowercase $\mathbf{a}$, and unbold letters A, a denote matrices, vectors, and scalar values, respectively. The symbols $\mathbb{E}(.)$, x^*, x^{T}, and x^{H} denote the expectation, conjugate, transpose, and Hermitian (conjugate transpose) of vector x, respectively. Additionally, the symbol $\mathbb{C}$ denotes the complex field, and $\mathbb{R}$ represents the real field. $\mathbb{C}^{M\times N}$ denotes the space of $M \times N$ complex valued matrices.

II. System Model

In this section, the system model of RSMA is presented, where both the transmission and reception architectures are described. After that, HARQ types are briefly introduced.

A. RSMA

Consider a single cell down-link network with one BS using N_t transmitting antennas serving K single-antenna users using

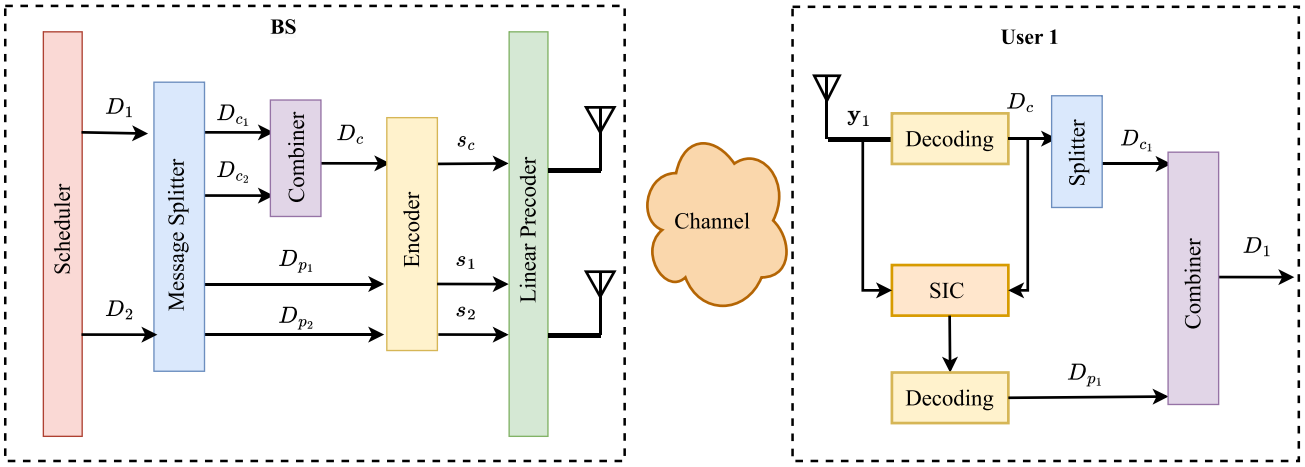

Fig. 1. RSMA system model.

RSMA as shown in Fig. 1.[1] In RSMA system, the message D_k of the kth user is split into a common part $D_{c,k}$ and a private part $D_{p,k}$. After that, the common parts of all the users $\{D_{c,1}, \ldots, D_{c,K}\}$ are jointly encoded into a common stream s_c, while the private parts $\{D_{p,1}, \ldots, D_{p,K}\}$ are, respectively, encoded into private streams $\{s_1, \ldots, s_K\}$. The data streams are denoted together as $\mathbf{s} = [s_c, s_1, \ldots, s_K]^T \in \mathbb{C}^{K+1}$, are linearly precoded before transmission using the precoder $\mathbf{P} = [\mathbf{p}_c, \mathbf{p}_1, \ldots, \mathbf{p}_K]$, where $\mathbf{p}_c, \mathbf{p}_k \in \mathbb{C}^{N_t \times 1}$ are the precoders of a common and the private streams, respectively. The power of the common stream is denoted as $P_c = \|\mathbf{p}_c\|_2$ and the transmitted signal of the kth user can be given as

$$\mathbf{s} = \sqrt{p_c} s_c + \sum_{k=1}^{K} \sqrt{p_k} s_k \quad \forall k \in K \tag{1}$$

where p_k and p_c are the transmitted powers of private and common submessages for kth user, respectively. Then, the received signal at kth user is shown

$$\mathbf{y}_k = \mathbf{h}_k^H \mathbf{s} + n_k \tag{2}$$

where $\mathbf{h}_k \in \mathbb{C}^{N_t \times 1}$ is the channel gain from BS to kth user and it is perfectly known at BS.[2] n_k is the AWGN at kth user with zero mean and variance σ_k^2 and it is represented as $n_k \sim \mathcal{CN}(0, \sigma_k^2)$. We assume that $\mathbf{s}$ follows $\mathbb{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$ and the total transmit power budget is P_t, i.e., $\text{tr}(\mathbf{P}\mathbf{P}^H) \leq P_t$. Each part of the splitted message has a maximum transmission power limit; for private messages P_k i.e., $\sum_{k=1}^{K} p_k \leq P_k$ and for common messages P_c i.e., $P_c \leq P_t$, therefore, $\sum_{k=1}^{K} (P_k + P_c) \leq P_t$.

According to the RSMA receiver, one of the main features is managing the interference where it is partially decoded and partially treated as noise. More specifically, the receiver of RSMA first detects the common stream and then applies SIC to be able to decode the corresponding private stream. After that, the receiver combines both decoded parts as shown in Fig. 1. Consequently, the signal to interference plus noise ratio (SINR) of both decoding parts, s_c and s_k, at the kth user are represented as

$$\gamma_c = \frac{\left|\mathbf{h}_k^H \mathbf{p}_c\right|^2}{\sum_{j \in K} \left|\mathbf{h}_k^H \mathbf{p}_j\right|^2 + \sigma_k^2} \quad \forall k \in K \tag{3}$$

$$\gamma_k = \frac{\left|\mathbf{h}_k^H \mathbf{p}_k\right|^2}{\sum_{j \in K, j \neq k} \left|\mathbf{h}_k^H \mathbf{p}_j\right|^2 + \sigma_k^2} \quad \forall k \in K. \tag{4}$$

B. HARQ

HARQ is a retransmission mechanism that employs channel coding across several rounds to retrieve messages that are incorrectly received. The two forms of HARQ are taken into consideration in this study; HARQ-chase combining (CC) and HARQ-incremental redundancy (IR). More specifically, in HARQ-CC, the generated codewords are the same across all the HARQ rounds. While, in HARQ-IR the generated codewords in various HARQ rounds are not similar. The main difference between the two lies in the way of handling the errors and retransmission of data.

Assuming that $\Lambda = (\gamma_1, \ldots, \gamma_M)$ is a total vector of SINR for the kth user where M is the maximum number of HARQ rounds, n is the packet length, and L is the information block length. Based on [14], the PER approximation for HARQ-CC and HARQ-IR is given below, respectively

$$\epsilon_{cc}\left(\left[\gamma_m\right]_0^{M-1}\right) \approx Q\left(\frac{n \log_2\left(1 + \sum_{m=0}^{M-1} \gamma_m\right) - L + \log_2(n)}{n\sqrt{V\left(\sum_{m=0}^{M-1} \gamma_m\right)}}\right) \tag{5}$$

$$\epsilon_{\text{ir}}\left(\left[\gamma_m\right]_0^{M-1}, \left[n_m\right]_0^{M-1}\right) \approx Q\left(\frac{\sum_{m=0}^{M-1} n_m \log_2(1 + \gamma_m) - L + \log_2\left(\sum_{m=0}^{M-1} n_m\right)}{\sqrt{\sum_{m=0}^{M-1} n_m V(\gamma_m)}}\right) \tag{6}$$

[1] It should be noted that although Fig. 1 shows two users using two transmit antennas, the associated system model figure may easily be expanded to include more users and antennas.

[2] Time-division duplexing is utilized to obtain CSI at BS by sending reference/pilot. It is used because of its simplicity and feature of channel reciprocity.

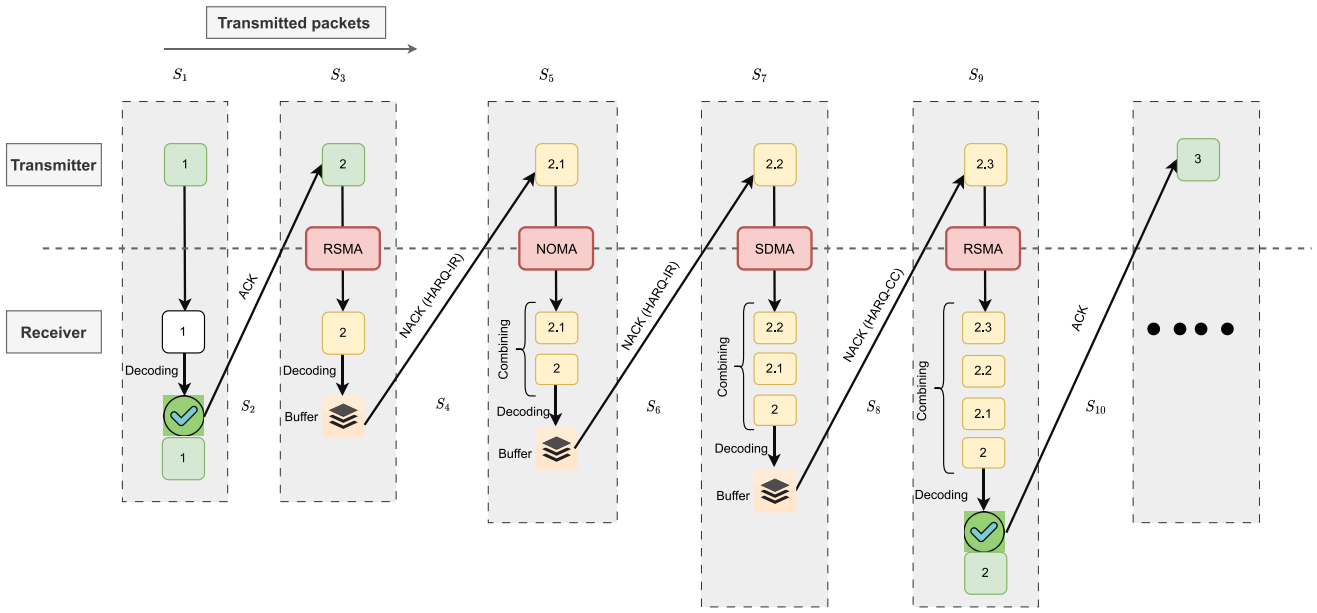

Fig. 2. Proposed HARQ-RSMA process.

where $Q(x)$ is the standard Q-function, and $V(\gamma_m) = (1-(1+\gamma_m)^{-2})\log_2^2(e)$ is a channel distortion and γ_m is the total SINR of the kth user which is as follows:

$$\gamma_m = \gamma_{k,m} + \gamma_{c,m} \tag{7}$$

where $\gamma_{c,m}$ and $\gamma_{k,m}$ represent the SINR for common and private at the mth round, respectively.

III. Proposed Retransmission Protocol

In this section, four aspects of the proposed HARQ-RSMA are presented. First, the proposed HARQ-RSMA is defined. Then, the retransmission process is explained and a new link adaptation approach is suggested. Finally, the pros and cons of the proposed approach are discussed.

A. Proposed HARQ-RSMA

Error detection, correction, retransmission, and combining are the steps of applying HARQ to wireless networks [18]. The proposed HARQ-RSMA structure is described as follows. The RSMA data frame, $D = [D_c, D_{p,1}, D_{p,2}, \ldots, D_{p,K}]$, is applied to an error detection encoder where l_{ed} bits are attached to D in order to discover errors on the receiver side. The output of the error detection encoder can be expressed as

$$\mathbf{c} = \left[\underbrace{c_0, c_1, c_2, \ldots, c_K}_{D_c, D_{p,1}, D_{p,2}, \ldots, D_{p,K}}, c_{k+1}, c_{k+2}, \ldots, c_{K+1+l_{\text{ed}}}\right]. \tag{8}$$

After that, each of the $\mathbf{c}$ sequence is applied to a channel encoder that attaches l_{ec} bits to $\mathbf{c}$ to form a code-word $\mathbf{u}$. The $(K+1)+l_{ed}$ is the number of bits applied to a channel encoder, and the parity bits can be extracted as $\mathbb{O} = [\mathbb{O}_0, \mathbb{O}_1, \mathbb{O}_2, \ldots, \mathbb{O}_{l_{ed}}]$. Then, a buffer of size, χ, is used to store the number of codewords to form a large block of bits $\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_\chi]$ which is interleaved to mitigate the effect of burst errors. After that, the interleaved output is modulated using a particular modulation scheme. The modulated code-word $s = [s_0, s_1, s_2, \ldots, s_{\mathbb{Q}}]$, where $\mathbb{Q}$ is the total number of modulated symbols in the sequence. It can be seen that the network has undergone several changes as a result of the adopted RSMA, including an increase in buffer size and the transmission of $\chi+1$ streams instead of χ streams according to the number of common streams in the RSMA network.

B. HARQ-RSMA Retransmission Process

The retransmission process of our suggested approach is depicted in Fig. 2 in ten distinct steps as $S_1, S_2, \ldots, S_{10}$. The following is the description of these steps.

1) First, the BS transmits a packet to the user equipment (UE), and the UE receives the packet and checks its FEC and reliability level, as shown by S_1 and S_3 blocks.
2) After that, the UE checks the system requirement, if it is matched, the received packet is decoded and UE transmits a unique instantaneous acknowledgment (ACK) to BS indicating the successful decoding, as shown by S_2. Otherwise, the UE stores the received packet and asks for the retransmission of the same packet by sending feedback as a negative acknowledgment (NACK) to the BS, as shown by S_4.
3) The BS retransmits based on the proposed strategies. In particular, the first retransmission strategy resends the common part of the data as an NOMA stream, since the error propagation problem comes from an error in decoding the common part, as shown by S_5.
4) Then, after checking the FEC and the required reliability of the system, the UE asks for another retransmission by sending a NACK to the BS, as shown by S_6. Here, the BS resends the private part as an SDMA stream which is the second strategy, as shown by S_7.

TABLE I
ADAPTIVITY FEATURE OF RSMA

	s_1	s_2	s_c
SDMA	D_1	D_2	$-$
NOMA	D_1	$-$	D_2
RSMA	$D_{p,1}$	$D_{p,2}$	$D_{c,1}, D_{c,2}$

5) The UE again checks the FEC bits, if the error is still not acceptable, the UE sends a NACK to the BS, as shown by S_8. In this case, the BS resends the whole data as an RSMA stream which is the third strategy, as shown by S_9. After that, the UE sends an ACK to the BS, as shown by S_{10}.

C. Link Adaptation Approach

In the literature, the link adaptation mechanism is an adaptation in modulation and coding approaches where both of them are changed based on the channel condition [19]. In this work, a new approach of link adaptation based on the multiaccess is proposed where two goals are achieved; retransmission and the way of retransmission (RSMA, NOMA, and SDMA). This is due to the flexibility that RSMA provides to networks. More specifically, NOMA is assigned to be with the first strategy to resend the common part, SDMA is assigned to the second strategy to resend the private part, and RSMA is assigned to the third strategy as shown in Fig. 2. Particularly, SDMA and NOMA are specific examples of RSMA, as seen in Table I.[3] SDMA is obtained from RSMA by forcing $\|\mathbf{p}_c\|^2 = 0$, therefore, D_k is directly encoded into s_k. Conversely, by encoding message D_2 entirely into s_c (i.e., $D_c = D_2$) and D_1 into s_1 while turning off $s_2(\|\mathbf{p}_2\|^2 = 0)$, we obtain NOMA [4].

D. Pros and Cons of the Proposed Approach

Proposing such a link adaptation approach adds another degree of freedom to the network along with modulation and coding, and this approach can effectively work with the existing link adaptation approach as illustrated in Fig. 3. Several advantages of adopting such a link adaptation approach can be summarized as follows.

1) Adopting such an approach can open several aspects of the research for beyond 5G networks, for example, proposing different transmission perspectives to be a part of a new adaptation scheme for future wireless networks.
2) Doing such adaptation based on the receiver side makes an adaptive and flexible combining process between the IR and CC approaches. In particular, the first and second strategies are combined as the IR approach while the third strategy is combined with the initial transmission of the packet as the CC approach which is shown in Fig. 2.
3) Using this kind of link adaptation opens several areas of research that are related to the coexistence approach between different beyond 5G networks for instance joint radar and communication, 5G new radio, and LTE.

[3]It should be noticed that although Table I only has two users, it is simple to expand the associated table to accommodate more.

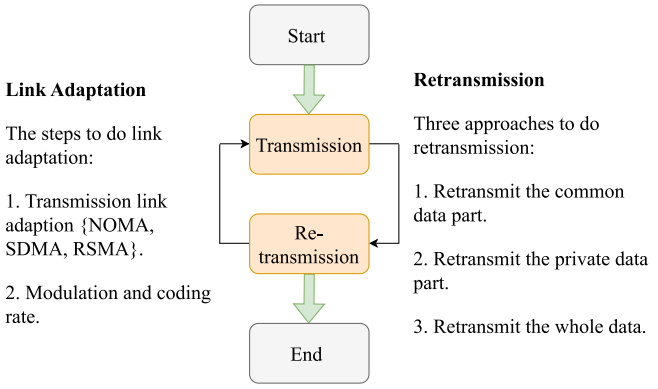

Fig. 3. Proposed link adaptation scheme.

Our proposed approach offers numerous advantages; however, several challenges require dedicated attention as future work. These challenges are vital to focus for further enhancement.

1) A pivotal challenge in our approach revolves around obtaining precise CSIT. Our proposed work assumes perfect CSIT, while the practical limitations of hardware introduce imperfections into the estimated CSIT. These inaccuracies lead to multiuser interference, imperfect SIC, and a reduction in the efficiency of link adaptation. Nonetheless, integrating machine learning and artificial intelligence methodologies holds the potential to refine the quality of CSIT estimation. By using such methods, we may improve channel estimates and precoder design, resulting in improved SIC implementation.
2) Although our approach provides flexibility in terms of using NOMA, SDMA, and RSMA during retransmission based on feedback signals, it may lead to increased signaling overhead compared to conventional HARQ methods. Particularly, traditional HARQ employs a single bit for feedback, our method utilizes multiple bits in a single time slot to indicate the specific type of retransmission (NOMA, SDMA, and RSMA).

IV. ANALYSIS OF THE PROPOSED APPROACH

In order to investigate the efficiency of the proposed design, two analyses are provided; the average number of retransmissions and EE. These analyses are studied for two reasons.

1) The average number of retransmissions is one of the most valuable constraints when it comes to scheduling techniques and minimizing the average age of information [20].
2) EE is a major issue while designing any network protocol since all users in a network have limited energy sources especially, IoT devices [21].

A. Average Number of Retransmissions

Considered Υ_m is the squared envelope of the channel gain during the mth retransmission round and the random variables $\Upsilon_m (m = 1, \ldots, M)$ are independent and identically distributed (i.i.d.). In the case of Rayleigh fading, the probability density

function (PDF) of Υ_m where $\bar{\upsilon}$ is the average received SINR, is given as

$$f_{\Upsilon_m}(\Upsilon_m) = \frac{e^{-\Upsilon_m/\bar{\upsilon}}}{\bar{\upsilon}}, \quad \Upsilon_m \geq 0. \tag{9}$$

The capacity of HARQ-RSMA after mth rounds of retransmission for common and private streams can be expressed as

$$C_m = \frac{1}{M}\sum_{m=1}^{M}\left[\log_2(1+\gamma_m\Upsilon_m)\right] \tag{10}$$

where γ_m is given in (7). The channel condition for RSMA plays a critical role in the required number of retransmissions, especially for common data packets since it causes an error propagation effect. More particular, a few HARQ cycles may be sufficient for error-free decoding if the channel state is good. On the other hand, if the channel faces deep fading, it could take numerous HARQ cycles to send a single data packet. Therefore, it is essential to define a transmission rate correctly for common and private parts.

The transmission rate of the first round for both common and private can be considered as $R_c = D_c/\mathbb{Q}$, and $R_k = D_{p,k}/\mathbb{Q}$, respectively, where $\mathbb{Q}$ is the number of transmitted symbols. More specifically, the required number of HARQ rounds to transmit a particular data packet either for common or private parts is determined by the transmission rate [22]. Consequently, if m HARQ rounds are used to deliver a particular data packet, $R_{c,m} = R_c/m$ is the transmission rate of a common stream, and $R_{k,m} = R_k/m$ is the transmission rate of the private streams. Then, the average transmission rate for common and private data packets is described as $\bar{R}_c = R_{c,m}/\mathbb{E}(T)$, $\bar{R}_k = R_{k,m}/\mathbb{E}(T)$, respectively, where the $\mathbb{E}(T)$ is the average number of retransmissions per data packet and T is a discrete random variable that represents the time between the initial sequence transmission and the successful decoding of the packet at the destination.

Assuming that T follows the Poisson distribution with an arrival rate equal to λ. In the absence of delay requirements in the network, the average rate is feasible. However, if the transmission fails to deliver a data packet after M (maximum number of rounds) HARQ rounds, the RSMA network would face an outage. In particular, if the total capacity of RSMA C_M is less than the total rate $R_c + R_k$, the network will face an outage. Then, the outage probability after M rounds is specified as

$$P_{\text{out},M}(R_c+R_k) = \mathbb{P}\left\{\left[\log_2(1+\gamma_m\Upsilon_m)\right] \leq R_c+R_k\right\}. \tag{11}$$

By considering a finite value of M and taking advantage of the outage probability in (11), the total average transmission rate of RSMA can be written as

$$\bar{R} = \frac{(R_c+R_k)\left(1-P_{\text{out},M}(R_c+R_k)\right)}{\mathbb{E}(T)}. \tag{12}$$

The $\mathbb{E}(T)$ can be given as

$$\mathbb{E}(T) = 1 + \sum_{m=1}^{M-1}\mathbb{P}\left(F^1,\ldots,F^m\right) \tag{13}$$

where F^m denotes the event of decoding the failure after m HARQ rounds either for common or private streams. Accordingly, three cases can be considered. If the data packet is successfully sent in the first round which is the correct event, $T = 1$. The probability of the correct event is denoted by $\mathbb{P}(S^1)$, whereas the probability of the failure event is as $\mathbb{P}(F^1)$. If the data packet transmission is unsuccessful in the first $M-1$ round and successful in the mth round, $T = M$. The probability of this occurrence is shown by $\mathbb{P}(F^1,\ldots,F^{m-1},S^m)$, while $\mathbb{P}(F^1,\ldots,F^{M-1},F^M)$ represents the likelihood that a transmission will fail after M rounds. If the data packet transmission fails in the first $M-1$ round then we experience either success or failure in the mth transmission round, $T = M$. Therefore, the probability $\mathbb{P}(F^1,\ldots,F^M)$ of a transmission failure after M rounds decreases as m increases and it is represented as follows:

$$\begin{aligned}\mathbb{P}\left(F^1,\ldots,F^M\right) &= P_{\text{out}}^M(R_c+R_k)\\ &= \mathbb{P}\left\{\sum_{m=1}^{M}\log_2(1+\gamma_m\Upsilon_m) \leq R_c+R_k\right\}\\ &= F_{\zeta_m}(R_c+R_k)\end{aligned} \tag{14}$$

where $F_{\zeta_m}(x)$ exemplifies the cumulative distribution function (CDF) of $\zeta_m = \sum_{m=1}^{M}\log_2(1+\gamma_m\Upsilon_m)$ where ζ_m is Ergodic capacity of HARQ-RSMA after m rounds for common and private streams. An approximate solution for the PDF and CDF of ζ_m are derived in the Appendix. The PDF of ζ_m can be written as

$$f_{\zeta_m}(x) \approx \frac{x^{\alpha_{\zeta_m}}}{\beta_{\zeta_m}^{(\alpha_{\zeta_m}+1)}\Gamma\left(\alpha_{\zeta_m}+1\right)}\exp\left(-\frac{x}{\beta_{\zeta_m}}\right) \tag{15}$$

where $\alpha_{\zeta_m} = ([\mathbb{E}(\zeta_m)]^2/\operatorname{Var}(\zeta_m)) - 1$, $\beta_{\zeta_m} = \operatorname{Var}(\zeta_m)\mathbb{E}(\zeta_m)$, and $\Gamma(\cdot)$ is the gamma function. $\operatorname{Var}(\zeta_m)$ and $\mathbb{E}(\zeta_m)$ are provided in the Appendix. The CDF of ζ_m can be written as

$$F_{\zeta_m}(x) = \frac{\alpha_{\zeta_m}!}{\Omega^{\alpha_{\zeta_m}+1}} - e^{-x\Omega}\sum_{l=0}^{\alpha_{\zeta_m}}\frac{\alpha_{\zeta_m}!}{l!}\frac{x^l}{\Omega^{\alpha_{\zeta_m}-l+1}} \tag{16}$$

where $\Omega = -1/\beta_{\zeta_m}$.

The Pollaczek–Khinchin equation is used to calculate the average waiting time for a data packet as follows [23]:

$$\mathbb{T} = \frac{\lambda\mathbb{E}(T^2)T_F}{2(1-\rho)} + \frac{T_F}{2} \tag{17}$$

where $\mathbb{E}(T^2)$ is the second-order moment of the total number of retransmissions, T_F is the frame duration, the utilization factor ρ, namely, the proportion of time during which the node is busy. The $\mathbb{E}(T^2)$ is expressed as [24]

$$\mathbb{E}\left(T^2\right) = 1 + \sum_{m=1}^{M-1}(2m+1)\mathbb{P}\left(F^1,\ldots,F^m\right). \tag{18}$$

The parameter ρ should satisfy the stability condition as follows:

$$\rho = \lambda\mathbb{E}(T)T_F < 1. \tag{19}$$

Consequently, the average waiting time is represented as

$$\mathbb{T} = \frac{T_F\lambda\left(1+\sum_{m=1}^{M-1}(2m+1)\mathbb{P}\left(F^1,\ldots,F^m\right)\right)}{2(1-\rho)} + \frac{T_F}{2}. \tag{20}$$

B. Energy Efficiency of Proposed Protocol

In the following, the EE analysis of the proposed HARQ-RSMA is investigated. The EE represented by η_{ee} is identified as follows [22]:

$$\eta_{ee} = \frac{\eta}{P} \tag{21}$$

where η is the throughput of RSMA network and P is the consumed power. P can be calculated by the multiplication of average of power consumed, $\bar{P}$, with the probability of successful transmission for each HARQ round. In particular, the probability of successful transmission and the probability of transmission failure in the first round can be represented as $\mathbb{P}(S^1)$ and $\mathbb{P}(F^1, \ldots, S^{M-1})$, respectively, until the $(M-2)$th round and success in the $(M-1)$th round. P can be shown for the proposed scheme as

$$P = \bar{P} \cdot \mathbb{P}\left(S^1\right) + 2\bar{P} \cdot \mathbb{P}\left(F^1, S^2\right) + \ldots + (M-1)\bar{P} \cdot \mathbb{P}\left(F^1, \ldots, S^{M-1}\right) + M\bar{P} \cdot \mathbb{P}\left(F^1, \ldots, F^{M-1}\right). \tag{22}$$

Therefore, (22) can be expressed as the product of the power consumed per HARQ round and the average number of retransmissions $\mathbb{E}[T]$. Then, P can be rewritten as

$$P = \bar{P} \cdot \mathbb{E}[T] = \bar{P}\left(1 + \sum_{m=1}^{M-1} \mathbb{P}\left(F^1, \ldots, F^m\right)\right). \tag{23}$$

The throughput of kth user of the RSMA network after applying the proposed HARQ protocol is expressed by the average sum rates of common and private parts as follows:

$$\mathcal{R}_c = \min\left\{\mathbb{E}\left[\ln\left(1 + \gamma_{c,m}\right)\right]\right\} \tag{24}$$

$$\mathcal{R}_k = \mathbb{E}\left[\ln\left(1 + \gamma_{\mathrm{k,m}}\right)\right] \tag{25}$$

then the achievable sum rates of HARQ-RSMA can be expressed as follows:

$$\eta = \mathcal{R}_c + \sum_{k=1}^{K} \mathcal{R}_k. \tag{26}$$

As a result, the EE of the proposed HARQ-RSMA can be written as

$$\eta_{ee} = \frac{\mathcal{R}_c + \sum_{k=1}^{K} \mathcal{R}_k}{\bar{P}\left(1 + \sum_{m=1}^{M-1} \mathbb{P}\left(F^1, \ldots, F^m\right)\right)}. \tag{27}$$

V. Analytical and Simulation Results

The analytical and simulation results are described in this section. Table II illustrates the simulation settings utilized to obtain the results, and the overall process of the proposed scheme is presented by Algorithm 1.

Fig. 4 (left) and (right) demonstrate the PER of the proposed HARQ-RSMA over AWGN and Rayleigh fading channels, respectively, taking into consideration two different coding schemes [convolutional codes and low-density parity check (LDPC)] with different data sizes ($L = 500$ and $L = 1500$). One can observe from Fig. 4 (left) that for the same length and different values of M, we need around 3-dB SNR to achieve 10^{-3} error rate. However, for the same number of retransmissions and different lengths, different error rates can be achieved at around 4-dB SNR. Also, from Fig. 4 (right), it is observed that the PER of the proposed scheme outperforms the conventional RSMA without HARQ by 7-dB SNR at 10^{-3} error rate (convolutional codes, $M = 3$, $L = 1500$) and the same trend is followed by LDPC (with $M = 3$, $L = 1500$). As a result, our proposed strategy can improve network reliability by 70% error reduction.

Moreover, our proposed approach performs better than both N-HARQ and orthogonal-HARQ (O-HARQ) that have been proposed in [14], where N-HARQ uses the nonorthogonality of the shared time slots to conduct the retransmissions

TABLE II
Simulation Parameters

Parameters	Assumptions
Modulation type	BPSK
Noise density (σ_k^2)	-80 dBmW/Hz
Coding scheme	Convolutional codes, LDPC
Maximum number of rounds (M)	3, 6
Data block length (L)	500, 1500 bit
Redundancy frame size	32 bit
Burst size (P_L)	2000 frames
Number of users (K)	2
Packet arrival rate (λ)	0.01
Frame duration (T_F)	1s
Code rate	1/3
Minimum acceptable error (E_{min})	10^{-6}
Channel type	AWGN, Rayleigh fading channel

Algorithm 1 Overall Algorithm of the Proposed Scheme

```
Input parameters: K; M; P_L; P;
Procedure:
for i=1:Length(SNR) do
  for k=1:K do
    for t=1:P_L do
      s = 0 {Initial buffer counter}
      PE = 0 {Initial packet error}
      while (PE > E_min & s < M) do
        if (s = 0) then
          {NACK is sent}
          {NOMA signal is transmitted}
        else if (s = 1) then
          {NACK is sent}
          {SDMA signal is transmitted}
        else if (s = 2) then
          {NACK is sent}
          {RSMA signal is transmitted}
        end if
        s = s + 1
      end while
    end for
    Compute: PE; E(T);
  end for
  Compute: η; η_ee;
end for
Outputs: PER; E(T); η_ee
```

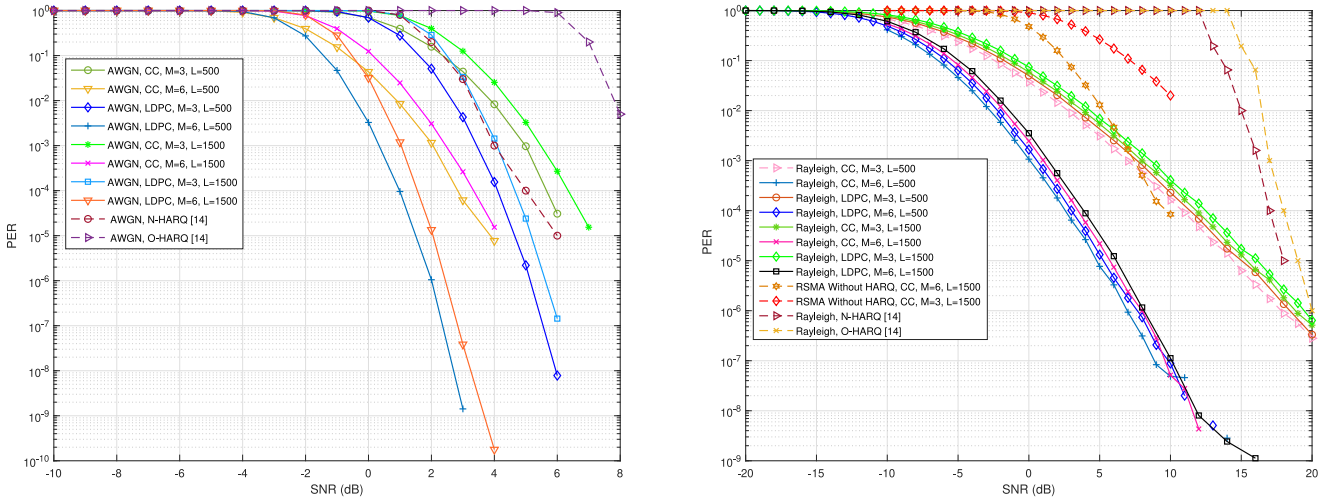

Fig. 4. Proposed HARQ-RSMA PER analysis over AWGN channel (left) and Rayleigh fading channel (right) considers convolutional and LDPC codes with varying transmission data sizes and maximum retransmissions within a HARQ round. Conv. C in the legend is short for the convolutional codes.

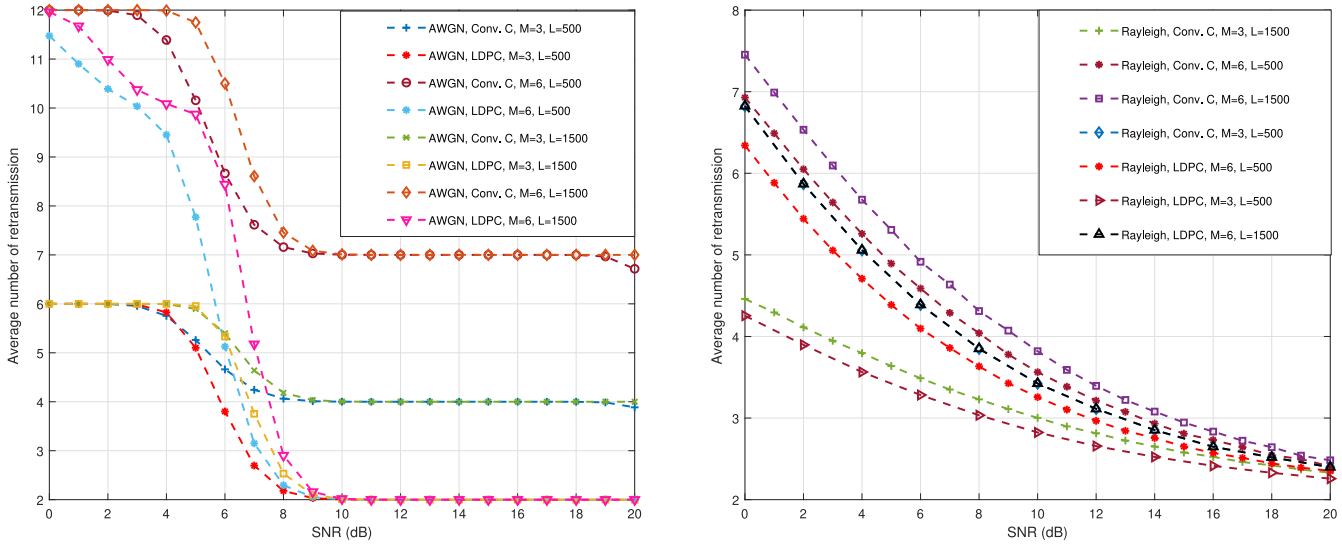

Fig. 5. Proposed HARQ-RSMA approach's average number of retransmissions over AWGN channel (left) and Rayleigh fading channel (right) takes into account convolutional and LDPC codes with varied transmission data sizes and maximum retransmissions within a HARQ round. Conv. C in the legend is short for the convolutional codes.

and O-HARQ uses a separate time slot for each packet to be retransmitted. O-HARQ represents the standard HARQ protocol.

The average number of retransmissions of the proposed HARQ-RSMA over AWGN and Rayleigh fading channels is shown in Fig. 5 (left) and (right), respectively. It can be clearly seen that the average number of HARQ-RSMA retransmissions over AWGN is restricted to low-to-medium SNR range and at higher values of SNR, it saturates. It is evident from the figures that at different sizes of data with the same number of retransmissions and coding schemes, the saturation level and average number of retransmissions are same. On the other hand, in Fig. 5 (right), the case of Rayleigh fading with convolutional codes, $M = 6$, and $L = 1500$, shows the maximum number of retransmissions that can be obtained. While, the case of Rayleigh fading with LDPC, $M = 6$, and $L = 500$, shows the minimum number of retransmissions that can be obtained. Notably, the average number of retransmissions using HARQ-RSMA decreases consistently as SNR increases, since an increase in the signal's power may reduce the chance of requesting a retransmission. Also, the type of FEC has a crucial effect on the average number of retransmissions and saturation level.

Fig. 6 (left) and (right) demonstrate the achievable sum rate of the proposed HARQ-RSMA over AWGN and Rayleigh fading channels, respectively. One can observe that the case of AWGN with convolutional codes, $M = 6$, and $L = 500$ and 1500, is giving the highest sum rate because we have a large data to be sent with convolutional codes which means a large number of transmissions and therefore the number of copies of this data increases at the receiver side. However, LDPC-based AWGN gives a slightly lower sum rate than the convolutional codes at the expense of lower average number of retransmissions of the same data as evident from Fig. 6

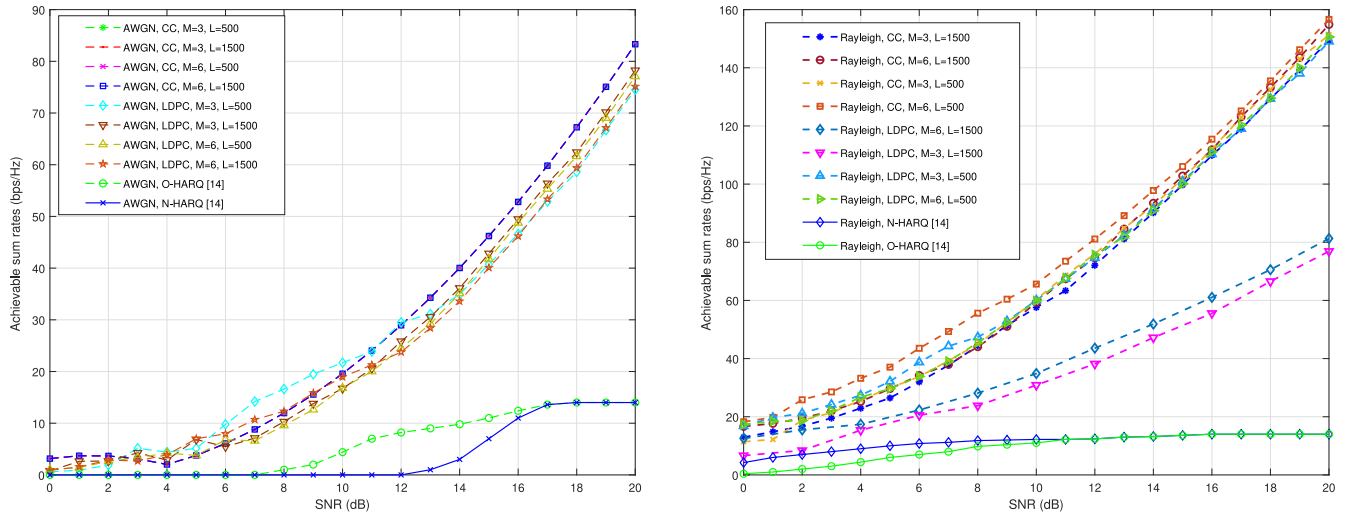

Fig. 6. Achievable sum rate of the proposed HARQ-RSMA approach over AWGN (left) and Rayleigh fading channel (right) takes into account convolutional and LDPC codes with different transmission data sizes and maximum retransmissions within a HARQ round. Conv. C in the legend is short for the convolutional codes.

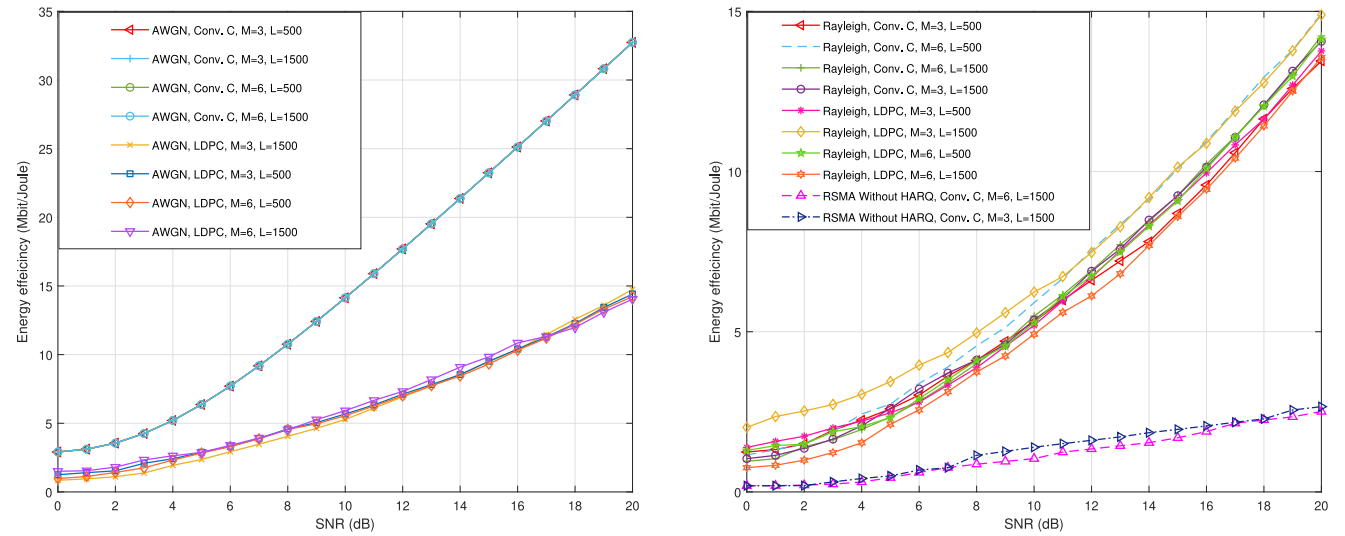

Fig. 7. EE versus SNR of the proposed HARQ-RSMA protocol over AWGN channel (left) and Rayleigh fading channel (right) takes into account convolutional and LDPC codes with different transmission data sizes and maximum retransmissions within a HARQ round. Conv. C in the legend is short for the convolutional codes.

(left). On the other hand, in Fig. 6 (right), it can be observed that the sum rate performance of the proposed HARQ-RSMA is greater than the conventional RSMA without HARQ, N-HARQ, and O-HARQ [14]. As a result, the sum rate of the proposed HARQ-RSMA continuously improves from low to high-SNR regimes because the proposed technique employs the link adaptation approach, which reduces the total number of retransmissions.

The EE of the proposed HARQ-RSMA over AWGN and Rayleigh fading channels is shown in Fig. 7 (left) and (right), respectively. It can be clearly seen from Fig. 7 (left), that convolutional codes based HARQ-RSMA have more EE compared with LDPC because they carry a large number of retransmissions using the same power. For instance, increasing the SNR (dB) from 10 to 18 will lead to an increase in EE by 1.5% with LDPC and convolutional codes. From Fig. 7 (right), it is obvious that the EE of the proposed HARQ-RSMA is larger than the conventional RSMA without HARQ. For illustration, at $SNR = 10$ dB, in the case of convolutional codes, $M = 6$, $L = 1500$, the proposed approach can achieve around 81% performance gain. It is worth noting that the EE of the proposed HARQ-RSMA improves steadily from lower to higher SNR regimes, however, in conventional RSMA without HARQ, it starts saturating at the higher SNR values. The proposed approach can enhance the power management capability which results in the improvement of EE.

A. Practical Considerations

To perform retransmission in RSMA networks, the proposed HARQ protocol has been designed based on three strategies where the BS can send three types of packets in a single HARQ round (RSMA, NOMA, and SDMA) and based on that a new link adaptation is suggested. As a result of providing

such flexibility based on multiple access methodologies, the complexity of the HARQ-RSMA transmitter grows where many factors need to be considered, such as power and rate allocations for each approach, resource management, and transmission signal formulation challenges. On the other hand, the receiver of the proposed HARQ protocol is able to carry out the SIC process and proper combining between different types of packets is performed based on IR and CC. Consequently, several delay factors come into the picture and have to be considered, namely, decoding delay (due to the SIC process), retransmission delay, queuing delay, processing delay, and feedback delay.

B. Computational Complexity Analysis

We investigated two forms of FEC codes while considering the computational complexity of the proposed HARQ-RSMA protocol: 1) convolutional code and 2) LDPC. When utilizing the convolutional code, the decoding methods of the Mth HARQ round need complexity in the order of $\mathcal{O}(3ML(\log ML))$. On the other hand, when considering the LDPC code, the decoding operations of the Mth transmission require complexities in the order of $\mathcal{O}((M+1)\log(ML))$, where $\mathcal{O}()$ refers to the big-$\mathcal{O}$ notation and L denotes the length of the data to be transmitted in each HARQ round.

VI. Conclusion and Future Work

In this article, we designed a novel HARQ protocol-based RSMA where three retransmission strategies are introduced. Furthermore, a link adaptable methodology based on RSMA, NOMA, and SDMA is provided. The proposed scheme outperforms the conventional RSMA networks in terms of reliability, acheivable sum rate, the average number of retransmissions, and EE due to its ability to select between NOMA, SDMA, and RSMA as retransmission strategies. Therefore, it is suitable to be applied to different networks, such as IoT networks, to get more reliable transmission and low-retransmission latency as it provides flexibility. Moreover, this work paves the way for many future directions.

1) This work can be substantial work for introducing RSMA to the 3GPP standards due to the low-complexity and compatibility benefits achieved by the proposed HARQ-RSMA design.
2) The suggested design might provide a new direction for medium access control layer and physical layer interactions which is one of the features of new generation of wireless communication.
3) Finally, investigating our proposed approach under other factors for instance retransmission delay and queuing delay can be considered as future work.

Appendix

Proof of (15)

We represent the total mutual information of the proposed HARQ-RSMA after m rounds by ζ_m where $\zeta_m = \sum_{m=1}^{M}[\log_2(1+\gamma_{t,m}\Upsilon_m)]$. An approximate solution for the PDF of ζ_m is provided in this Appendix. By using the Laguerre series [25], the PDF of the ζ_m can be represented as

$$f_{\zeta_m}(x) \approx \frac{x^{\alpha_{\zeta_m}}}{\beta_{\zeta_m}^{(\alpha_{\zeta_m}+1)}\Gamma(\alpha_{\zeta_m}+1)}\exp\left(-\frac{x}{\beta_{\zeta_m}}\right) \tag{28}$$

where $\alpha_{\zeta_m} = ([\mathbb{E}(\zeta_m)]^2/\operatorname{Var}(\zeta_m)) - 1$, $\beta_{\zeta_m} = \operatorname{Var}(\zeta_m)/\mathbb{E}(\zeta_m)$, and $\Gamma(\cdot)$ is the gamma function. To determine the parameters α_{ζ_m} and β_{ζ_m}, the mean value and the variance of the process ζ_m have to be defined. Assuming that the average SINR per receive antenna is $\bar{\upsilon}$, the PDF of the MRC SINRs at the MIMO system's output can be represented as follows:

$$f_{\Upsilon_m}(\Upsilon_m) = \sum_{b=1}^{B}\sum_{l=r-B}^{(r+B-2b)\ b}\frac{a'_{b,l}}{l!}\left(\frac{b}{\bar{\upsilon}}\right)^{l+1}\Upsilon_m^l e^{-\frac{b}{\bar{\upsilon}}\Upsilon_m} \tag{29}$$

where $a'_{b,l} \triangleq a_{b,l}K_{B,r}l!/b^{l+1}$. Observing that $f_{\Upsilon_m}(\Upsilon_m)$ is a weighted sum of elementary Gamma PDFs with parameter $l+1$ and mean $(l+1)\bar{\upsilon}/b$. $K_{B,r} = (\prod_{i=1}^{B}(B-i)!\prod_{i=1}^{B}(r-i)!)^{-1}$ is a normalizing constant [26]. The expected value of the ζ_m can be written as follows taking into account the processes Υ_m are i.i.d.

$$\begin{aligned}\mathbb{E}(\zeta_m) &= \sum_{m=1}^{M}\mathbb{E}\{\log_2(1+\gamma_m\Upsilon_m)\}\\ &= M\mathbb{E}\{\log_2(1+\gamma_m\Upsilon_m)\}.\end{aligned} \tag{30}$$

The term $\mathbb{E}\{\log_2(1+\gamma_m\Upsilon_m)\}$ is equal to the Ergodic capacity C of the Rayleigh fading channel which can be written as

$$\begin{aligned}C &= \int_0^\infty \log_2(1+\gamma_m\Upsilon_m)f_{\Upsilon_m}(\Upsilon_m)d\Upsilon_m\\ &= \int_0^\infty \log_2(1+\gamma_m\Upsilon_m)\sum_{b=1}^{B}\sum_{l=r-B}^{(r+B-2b)\ b}\frac{a'_{b,l}}{l!}\left(\frac{b}{\bar{\upsilon}}\right)^{l+1}\\ &\quad \Upsilon_m^l e^{-\frac{b}{\bar{\upsilon}}\Upsilon_m}d\Upsilon_m\\ &= Z\int_0^\infty \ln(1+\gamma_m\Upsilon_m)\Upsilon_m^l e^{-\frac{b}{\bar{\upsilon}}\Upsilon_m}d\Upsilon_m\end{aligned} \tag{31}$$

where Z is represented as follows:

$$Z = \log_2(e)\sum_{b=1}^{B}\sum_{l=r-B}^{(r+B-2b)\ b}\frac{a'_{b,l}}{l!}\left(\frac{b}{\bar{\upsilon}}\right)^{l+1}. \tag{32}$$

The integral in (31) is evaluated using partial integration as follows, $u = \ln(1+t)$ and $du = [(dt)/(1+t)]$ are related to the first part, and $dv = t^{n-1}e^{-\mu t}dt$ is related to the second part. Then

$$\int_0^{+\infty} udv = \lim_{t\to+\infty}(uv) - \lim_{t\to 0}(uv) - \int_0^{+\infty} vdu. \tag{33}$$

Performing $n-1$ successive integration by parts yields [27] [p. 112 and eq. (2.321.2)]

$$v = -e^{-\mu t}\sum_{k=1}^{n}\frac{(n-1)!}{(n-k)!}\frac{t^{n-k}}{\mu^k}. \tag{34}$$

Hence, the integral can be written as

$$I_n(\mu) = \sum_{k=1}^{n}\frac{(n-1)!}{\mu^k(n-k)!}\int_0^{+\infty}\frac{t^{n-k}e^{-\mu t}}{1+t}dt. \tag{35}$$

The integral $I_n(\mu)$ can be simplified in a short form using [27] [p. 366 and eq. (3.383.10)], giving

$$I_n(\mu) = (n-1)!e^{\mu}\sum_{k=1}^{n}\frac{\Gamma(-n+k,\mu)}{\mu^k} \tag{36}$$

where $\Gamma(\cdot,\cdot)$ is the complementary incomplete gamma function represented by [27] [p. 949 and eq. (8.350.2)]

$$\Gamma(\alpha,x) = \int_x^{+\infty} t^{\alpha-1}e^{-t}dt. \tag{37}$$

As a result, the C can be written as following:

$$C = Z(n-1)!e^{\mu}\sum_{k=1}^{n}\frac{\Gamma(-n+k,\mu)}{\mu^k}. \tag{38}$$

The variance of ζ_m can be expressed as

$$\begin{aligned}\mathrm{Var}(\zeta_m) &= \sum_{m=1}^{M}\mathrm{Var}\left(\left\{\log_2(1+\gamma_m\Upsilon_m)\right\}\right)\\ &= M\mathbb{E}\left\{\left(\log_2(1+\gamma_m\Upsilon_m)\right)^2\right\}\\ &\quad - M\mathbb{E}\left\{\log_2(1+\gamma_m\Upsilon_m)\right\}^2.\end{aligned} \tag{39}$$

The second part of the variance represents the mean of ζ_m that is equal to (38). Therefore, we have to solve the first part as follows:

$$\begin{aligned}&\mathbb{E}\left\{\left[\log_2(1+\gamma_m\Upsilon_m)\right]^2\right\}\\ &\quad= \int_0^{\infty}\left[\log_2(1+\gamma_m\Upsilon_m)\right]^2 f_{\Upsilon_m}(\Upsilon_m)d\Upsilon_m\\ &\quad= \left(\log_2(e)\right)^2\int_0^{\infty}\left[\ln(1+\gamma_m\Upsilon_m)\right]^2 f_{\Upsilon_m}(\Upsilon_m)d\Upsilon_m.\end{aligned} \tag{40}$$

The integral of the previous equation can be solved using the integration of parts. Assuming, $u = \ln(1+\gamma_m\Upsilon_m)^2$, thus, $du = 2\ln(1+\gamma_m\Upsilon_m)(\gamma_m d\Upsilon_m)/(1+\gamma_m\Upsilon_m)$. Then, let $dv = f_{\Upsilon_m}(\Upsilon_m)$, therefore, $v = 1$. So, the integral follows:

$$\int_0^{+\infty} udv = -\int_0^{+\infty} 2\ln(1+\gamma_m\Upsilon_m)(\gamma_m/(1+\gamma_m\Upsilon_m))d\Upsilon_m. \tag{41}$$

This integral can be solved using [27] [p. 14 and eq. (4.293)] as

$$\int_0^{\infty}\frac{x\ln(1+x)}{(1+x)}dx = \mathrm{B}(2,-1)[\psi(1)-\psi(-1)] \tag{42}$$

where

$$\mathrm{B}(x,y) = \frac{1}{y}\sum_{n=0}^{\infty}(-1)^n y\frac{(y-1)\cdots(y-n)}{n!(x+n)} \tag{43}$$

and

$$\psi(x) = \ln x - \sum_{k=0}^{\infty}\left[\frac{1}{x+k} - \ln\left(1+\frac{1}{x+k}\right)\right]. \tag{44}$$

Hence, the variance in (40) is represented by as

$$\mathrm{Var}(\zeta_m) = \log_2(e)^2\mathrm{B}(2,-1)[\psi(1)-\psi(-1)]. \tag{45}$$

After obtaining the $\mathbb{E}(\zeta_m)$ and $\mathrm{Var}(\zeta_m)$, the parameters α_{ζ_m} and β_{ζ_m} can be calculated using the expression of the PDF in (29). The CDF of ζ_m can be determined using

$$F_{\zeta_m}(x) = \int_0^x p_{\zeta_m}(z)dz. \tag{46}$$

This integral can be solved using [27] [p. 389 and eq. (3.351.1)], as follows:

$$\int_0^x z^{\alpha_{\zeta_m}}e^{-\Omega z}dz = \frac{\alpha_{\zeta_m}!}{\Omega^{\alpha_{\zeta_m}+1}} - e^{-x\Omega}\sum_{k=0}^{\alpha_{\zeta_m}}\frac{\alpha_{\zeta_m}!}{k!}\frac{x^k}{\Omega^{\alpha_{\zeta_m}-k+1}} \tag{47}$$

where Ω is

$$\Omega = -\frac{1}{\beta_{\zeta_m}}. \tag{48}$$

ACKNOWLEDGMENT

The authors are thankful to Dr. Haji M. Furqan from VESTEL for his constructive comments.

REFERENCES

[1] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*, vol. 159, Cham, Switzerland: Springer, 2019.

[2] A. Anwar, B.-C. Seet, M. A. Hasan, and X. J. Li, "A survey on application of non-orthogonal multiple access to different wireless networks," *Electronics*, vol. 8, no. 11, p. 1355, 2019.

[3] H. V. Nguyen, H. M. Kim, G.-M. Kang, K.-H. Nguyen, V.-P. Bui, and O.-S. Shin, "A survey on non-orthogonal multiple access: From the perspective of spectral efficiency and energy efficiency," *Energies*, vol. 13, no. 16, p. 4106, 2020.

[4] B. Clerckx, Y. Mao, R. Schober, and H. V. Poor, "Rate-splitting unifying SDMA, OMA, NOMA, and multicasting in MISO broadcast channel: A simple two-user rate analysis," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 349–353, Mar. 2019.

[5] Y. Mao, B. Clerckx, and V. O. Li, "Rate-splitting multiple access for downlink communication systems: Bridging, generalizing, and outperforming SDMA and NOMA," EURASIP *J. Wireless Commun. Netw.*, vol. 2018, pp. 1–54, 2018. [Online]. Available: https://doi.org/10.1186/s13638-018-104-7

[6] S. Gamal, M. Rihan, S. Hussin, A. Zaghloul, and A. A. Salem, "Multiple access in cognitive radio networks: From orthogonal and non-orthogonal to rate-splitting," *IEEE Access*, vol. 9, pp. 95569–95584, 2021.

[7] A. Ahmed, A. Al-Dweik, Y. Iraqi, H. Mukhtar, M. Naeem, and E. Hossain, "Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 4, pp. 2711–2752, 4th Quart., 2021.

[8] "Evolved universal terrestrial radio access (E-UTRA); medium access control (MAC) protocol specification; (Release 12), Version 12.5.0," 3GPP, Sophia, Antipolis, France, Std. TS 36.321, 2021. [Online]. Available: http://www.3gpp.org/DynaReport/36321.htm

[9] X. Leturc, C. J. Le Martret, and P. Ciblat, "Energy efficient resource allocation for HARQ with statistical CSI in multiuser ad hoc networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2017, pp. 1–6.

[10] Z. Shi, C. Zhang, Y. Fu, H. Wang, G. Yang, and S. Ma, "Achievable diversity order of HARQ-aided downlink NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 471–487, Jan. 2019.

[11] F. Ghanami, G. A. Hodtani, B. Vucetic, and M. Shirvanimoghaddam, "Performance analysis and optimization of NOMA with HARQ for short packet communications in massive IoT," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4736–4748, Mar. 2020.

[12] Z. Xiang, W. Yang, G. Pan, Y. Cai, Y. Song, and Y. Zou, "Secure transmission in HARQ-assisted non-orthogonal multiple access networks," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2171–2182, 2020.

[13] Z. Shi, S. Ma, H. ElSawy, G. Yang, and M.-S. Alouini, "Cooperative HARQ-assisted NOMA scheme in large-scale D2D networks," *IEEE Trans. Commun.*, vol. 66, no. 9, pp. 4286–4302, Sep. 2018.

[14] F. Nadeem, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Nonorthogonal HARQ for URLLC: Design and analysis," *IEEE Internet Things J.*, vol. 8, no. 24, pp. 17596–17610, Dec. 2021.

[15] J. Xu, O. Dizdar, and B. Clerckx, "Rate-splitting multiple access for short-packet uplink communications: A finite blocklength analysis," *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 517–521, Feb. 2023.
[16] Y. Xu, Y. Mao, O. Dizdar, and B. Clerckx, "Max-min fairness of rate-splitting multiple access with finite blocklength communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 5, pp. 6816–6821, May 2023.
[17] Y. Mao, O. Dizdar, B. Clerckx, R. Schober, P. Popovski, and H. V. Poor, "Rate-splitting multiple access: Fundamentals, survey, and future research trends," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 4, pp. 2073–2126, 4th Quart., 2022.
[18] H. Mukhtar, A. Al-Dweik, M. Al-Mualla, and A. Shami, "Low complexity power optimization algorithm for multimedia transmission over wireless networks," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1, pp. 113–124, Feb. 2015.
[19] P. H. Tan, Y. Wu, and S. Sun, "Link adaptation based on adaptive modulation and coding for multiple-antenna OFDM system," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1599–1606, Oct. 2008.
[20] E. T. Ceran, D. Gündüz, and A. György, "A reinforcement learning approach to age of information in multi-user networks with HARQ," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1412–1426, May 2021.
[21] Z. Wang, R. Liu, Q. Liu, J. S. Thompson, and M. Kadoch, "Energy-efficient data collection and device positioning in UAV-assisted IoT," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1122–1139, Feb. 2019.
[22] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1971–1988, Jul. 2001.
[23] W. Chan, T. Lu, and R. Chen, "Pollaezek-Khinchin formula for the M/G/1 queue in discrete time with vacations," *IEE Proc.-Comput. Digit. Techn.*, vol. 144, no. 4, pp. 222–226, Jul. 1997.
[24] A. Chelli, E. Zedini, M.-S. Alouini, J. R. Barry, and M. Pätzold, "Performance and delay analysis of hybrid ARQ with incremental redundancy over double rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6245–6258, Nov. 2014.
[25] S. Primak, V. Kontorovich, and V. Lyandres, *Stochastic Methods and Their Applications to Communications: Stochastic Differential Equations Approach*. Chichester, NH, USA: Wiley, 2005.
[26] A. Maaref and S. Aissa, "Closed-form expressions for the outage and Ergodic Shannon capacity of MIMO MRC systems," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1092–1095, Jul. 2005.
[27] I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Cambridge, MA, USA: Academic press, 2007.

Shaima Abidrabbu received the B.S. degree (with Distinguished Ratings) in communication and software engineering from Al-Balqa Applied University, Salt, Jordan, in 2015, and the M.Sc. degree (with Distinguished Ratings) in wireless communication engineering from Jordan University of Science and Technology, Irbid, Jordan, in 2019. She is currently pursuing the Ph.D. degree with the Department of Engineering and Natural Sciences, Istanbul Medipol University, Istanbul, Turkey.

She is also a Specialist with the Department of Intellectual Property Rights and License Agreements, Vestel Electronics, Manisa, Turkey. Her research interests lie in the field of wireless communications with emphasis on cognitive radio networks, power-domain NOMA, game theory, Wi-Fi networks, and physical-layer security.

Sawaira Rafaqat Ali received the bachelor's degree in telecommunication engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2021. She is currently pursuing the master's degree in electrical electronics and cyber system engineering with Istanbul Medipol University, Istanbul, Turkey.

Her research is centered on the domain of wireless communication with a primary focus on multiple access techniques for future wireless networks.

Hüseyin Arslan (Fellow, IEEE) received the B.S. degree from the Middle East Technical University, Ankara, Turkey, in 1992, and the M.S. and Ph.D. degrees from Southern Methodist University, Dallas, TX, USA, in 1994 and 1998, respectively.

From January 1998 to August 2002, he was with the Research Group of Ericsson, where he was involved with several projects related to 2G and 3G wireless communication systems. From August 2002 to 2022, he was a Professor with the Electrical Engineering Department, University of South Florida, Tampa, FL, USA. In December 2013, he joined Istanbul Medipol University, Istanbul, Turkey, to found the Engineering College, where he has been working as the Dean of the School of Engineering and Natural Sciences. In addition, he has worked as a part-time Consultant for various companies and institutions, including Anritsu Company, Atsugi, Japan, and The Scientific and Technological Research Council of Turkey. He conducts research in wireless systems, with emphasis on the physical and medium access layers of communications. He has been collaborating extensively with key national and international industrial partners and his research has generated significant interest in companies, such as InterDigital, Anritsu, NTT DoCoMo, Raytheon, Honeywell, and Keysight Technologies. Collaborations and feedback from industry partners have significantly influenced his research. In addition to his research activities, he has also contributed to wireless communication education. He has integrated the outcomes of his research into education which lead him to develop a number of courses with the University of South Florida. He has developed a unique "Wireless Systems Laboratory" course (funded by the National Science Foundation and Keysight technologies) where he was able to teach not only the theory but also the practical aspects of wireless communication system with the most contemporary test and measurement equipment. His current research interests are on 5G and beyond radio access technologies, physical-layer security, interference management (avoidance, awareness, and cancelation), cognitive radio, multicarrier wireless technologies (beyond OFDM), dynamic spectrum access, co-existence issues, nonterrestrial communications (high-altitude platforms), joint radar (sensing), and communication designs.

Dr. Arslan has served as a General Chair, a Technical Program Committee Chair, a Session and Symposium Organizer, a Workshop Chair, and a Technical Program Committee Member in several IEEE conferences. He is a member of the editorial board for the IEEE SURVEYS AND TUTORIALS AND THE SENSORS JOURNAL. He has also served as a member of the editorial board for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, and several other scholarly journals by Elsevier, Hindawi, and Wiley Publishing.