Received February 2, 2021, accepted February 10, 2021, date of publication February 12, 2021, date of current version February 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3059069

Analysis of NOMA-Based Retransmission Schemes for Factory Automation Applications

ENEKO IRADIER[1], (Graduate Student Member, IEEE),
LORENZO FANARI[1], (Graduate Student Member, IEEE),
IÑIGO BILBAO[1], (Graduate Student Member, IEEE),
JON MONTALBAN[1], (Senior Member, IEEE), PABLO ANGUEIRA[1], (Senior Member, IEEE),
OSCAR SEIJO[2], (Member, IEEE), AND IÑAKI VAL[2], (Senior Member, IEEE)
[1]Department of Communication Engineering, University of the Basque Country (UPV/EHU), 48013 Bilbao, Spain
[2]IKERLAN Technology Research Centre, Basque Research and Technology Alliance (BRTA), 20500 Mondragón, Spain

Corresponding author: Eneko Iradier (eneko.iradier@ehu.eus)

This work was supported in part by the Basque Government under Grant IT1234-19, in part by the PREDOC under Grant PRE_2020_2_0105, and in part by the Spanish Government through project PHANTOM (MCIU/AEI/FEDER, UE) under Grant RTI2018-099162-B-I00.

ABSTRACT New use cases and applications in factory automation scenarios impose demanding requirements for traditional industrial communications. In particular, latency and reliability are considered as some of the most representative Key Performance Indicators (KPI) that limit the technological choices addressing wireless communications. Indeed, there is a considerable research effort ongoing in the area of wireless systems, not only from academia, but also from companies, towards novel solutions that fit Industry 4.0 KPIs. A major limitation for traditional wireless architectures is related to the harsh nature of the industrial propagation channel. Accordingly, this paper addresses these challenges by studying the reliability and latency performance of the joint use of different retransmission schemes in combination with Non-Orthogonal Multiple Access (NOMA) techniques. Two general retransmission schemes have been tested: time-based and spatial diversity-based retransmissions. An adaptive injection level NOMA solution has been combined with the retransmission schemes to improve the reliability of critical information. In all cases, a particular set of simulations has been carried out varying the main parameters, such as modulation, code rate and the injection level. Moreover, the impact of the number of transmitters in relation to the communication reliability has been analyzed. Results show that spatial diversity-based retransmissions overcome considerably the reliability obtained with time-domain retransmissions while maintaining assumable latency rates.

INDEX TERMS 802.11, factory automation, industry 4.0, LDM, NOMA, P-NOMA, retransmissions, spatial diversity, wireless communications.

I. INTRODUCTION

Wireless communications are considered one of the most challenging and promising research areas of the so-called Industry 4.0 [1], [2]. In fact, the high costs and the lack of scalability and mobility that traditional automated processes composed by wired control systems have to assume make this solution less efficient than wireless systems. Moreover, the deployment of wireless systems provides the opportunity to enhance traditional systems by introducing new capabilities.

The associate editor coordinating the review of this manuscript and approving it for publication was Biju Issac.

However, the main challenge is the deployment scenario. Industrial environments entail strict requirements that are still hard to meet by wireless technologies, especially in the Factory Automation (FA) cases. For instance, very low latency values (0.25-10 *ms*) and ultra-high reliability (error rates below to 10^{-9}) are required [3], [4]. Determinism is another mandatory condition that is difficult to achieve with current wireless standards. In recent literature, in order to enable a massive deployment of industrial wireless networks, several technologies and standards have been proposed: IEEE 802.11 [5], [6], Bluetooth [7], 802.15.4 [8], LTE [9], and 5G [10], [11]. However, the latest versions of Wi-Fi and 5G stand out among all of them for a considerable improvement in critical applications. On the one hand, the next generation Wi-Fi standard (i.e., IEEE 802.11be) will offer, among other things, full-duplex multi-channel communications, which makes it easier to meet the reliability and latency requirements established by the industry [12]. On the other hand, 5G also presents conditions to meet industrial requirements due to its PHY/MAC layer with a wide range of possible configuration parameters. Among them, 5G allows to use small OFDM symbols (i.e., below 10 μs) with very robust modulation and code rates [13].

Since current standard wireless technologies do not cope with the demanding requirements of industrial applications and future standards such as 5G or IEEE 802.11be are still far from industrial deployment, some researchers have proposed proprietary solutions in order to get close to those barriers. For example, in [14], [15], Luvisotto *et al.* proposed the joint optimization of the PHY and MAC layers, improving the PHY reliability in several orders of magnitude and reducing the system latency by introducing a reliable MAC layer. In [5], [16], the authors follow the same joint PHY and MAC optimization approach in the solution referred to as SHARP (Synchronous and Hybrid Architecture for Real-time Performance). SHARP is based on the 802.11g standard and adds determinism to the PHY and MAC layers. In order to obtain competitive reliability and latency values, on the one hand, PHY frame aggregation techniques are implemented. On the other hand, a TDMA-based MAC structure with controlled time-division retransmissions is designed to guarantee a deterministic behavior.

A different strategy, based on improving the architecture of the access network is explored in [17], where several MAC level techniques in combination with the Wi-Fi PHY layer in challenging wireless channels are tested. Introducing multiple transmitters provided a reliability gain range of 8-12 dB. Specifically, the tests evaluated the use of two, three and four transmitters and they showed a non-linear behavior of the reliability increase while increasing the number of transmitters. Recently, in [18], the performance of spatial diversity with multiple redundant transmitters for mobile industrial communications is experimentally evaluated in a testbed with mobile robots. Authors implemented diversity using the Multipath TCP (MPTCP) protocol, which allows establishing different TCP connections between a transmitter and a receiver using multiple disjoint paths. The results indicate that redundant retransmissions improve reliability without compromising latency.

In [19], the use of Non-Orthogonal Multiple Access (NOMA) in combination with the IEEE 802.11n standard to meet the strict industrial requirements is considered. These results show that NOMA has the potential for enabling highly robust communications for industrial environments. However, the combination of NOMA and retransmission techniques has not been evaluated as a potential way to get closer to the performance requirements of current industry communications.

Therefore, in this paper, we propose and evaluate the joint use of NOMA and different diversity schemes to enhance the latency and reliability in industrial wireless networks. The diversity schemes are based on temporal and spatial diversity-based retransmissions using NOMA. Besides, some guidelines and recommendations are presented for real hardware implementation. In summary, the technical contributions of this paper include:

1) A comprehensive analysis of the PHY performance of a combined 802.11n and NOMA transceiver.
2) An analysis of the Injection Level (IL) capability to enhance the reliability.
3) A detailed analysis and performance evaluation of time-domain retransmissions.
4) Proposal and evaluation of combined multiple transmitter retransmissions with NOMA techniques.
5) Proposal and evaluation of a MAC layer that includes time, spatial and layer division.

The rest of the paper is organized as follows. The next section describes the related work, including the basis of NOMA techniques and state-of-the-art of retransmission and diversity schemes. Section III is focused on the presentation of the industrial use case and the simulation methodology. In Section IV, the PHY/MAC layers of the NOMA-based 802.11n transceiver are presented and evaluated. Then, in Section V, time-domain retransmissions are tested. Section VI is focused on the design and evaluation of MAC schemes that include multiple synchronized transmitters. Afterward, in Section VII a MAC layer that includes time, space and power multiplexing is designed and evaluated. Then, the latency and reliability results of the three different solutions are discussed and compared in Section VIII. Finally, Section IX contains the conclusions of the article.

II. RELATED WORK

This section contains, first, an overview of NOMA techniques and, then, a state-of-the-art of retransmission schemes including spatial diversity.

A. NOMA: GENERAL CONCEPTS AND INDUSTRIAL APPLICABILITY USING 802.11

NOMA represents a set of different medium access techniques where the receivers use the resource (either space, frequency, and time) in a non-orthogonal way. According to [20], NOMA techniques can be divided into two main families: code-domain NOMA (C-NOMA) [21], [22] and power-domain NOMA (P-NOMA) [23], [24]. In particular, although both NOMA families have been repeatedly proposed as alternatives to be included in 3GPP standards [25], P-NOMA techniques have shown a better complexity/performance tradeoff than C-NOMA [20]. During the rest of the paper, NOMA will refer to P-NOMA.

NOMA consists of different signals organized in several layers, where each layer takes some part of the transmitted total power. Each layer of the NOMA ensemble can be independently configured in order to address different reception targets. Each NOMA configuration depends on the modulation and the coding choices assigned to each layer, and on the injection level (Δ, measured in dB), defined as the power splitting ratio among layers.

The main benefit of NOMA in comparison with classical TDM/FDM systems is the increase in spectral efficiency [26].

In [19], the first approach of a NOMA-based 802.11n system was proposed for FA environments. In this work, the communication architecture (PHY/MAC) of the 802.11n standard to include NOMA features was redesigned. Then, the proposed architecture was evaluated and the results showed a considerable better performance than the 802.11n standard PHY on several industrial use cases. In [27], a specific application of the NOMA-based solution was proposed for an industrial multimedia content broadcasting environment.

However, the reliability obtained in the previous works is still not enough for the most challenging industrial applications. Following the approach taken in previous works [14], advances in PHY reliability can be complemented with different MAC level techniques. Specifically, it is expected that MAC layer enhancements can improve the 10^{-4} Packet Error Rate (PER) provided by the physical layer down to 10^{-8}. That is why, in this paper, different MAC level techniques are proposed and implemented in combination with the NOMA-based 802.11n prototype.

B. RETRANSMISSIONS AND DIVERSITY SCHEMES

In communication systems, especially in industrial environments, the design of a robust PHY layer goes together with an efficient MAC layer. Among others, the main functionalities of the MAC layer are: the management of the transmission/reception instants based on the traffic requirements and on the wireless medium conditions, the configuration of the PHY layer, and the implementation of diversity mechanisms. One of the most widespread techniques is the use of retransmissions [28], with a variety of different configurations.

First, the authors in [17] evaluate time-domain retransmissions over 802.11n jointly with a TDMA scheme for different time-spacings between the first transmissions and the concurrent retransmissions. They determine that the effectiveness of these methods and the latency associated with them are closely related, since the more time that passes between transmission and retransmission, the greater the probability of success, but the greater the latency. In [29], the authors point out that wireless interferences can appear with different characteristics of intensity and duration. Therefore, they propose a MAC layer with temporal diversity that adapts the retransmissions to the characteristics of the interferences. The results show that their proposal can be applied for different types of information, meeting the requirements of real-time communications.

The retransmission of erroneous packets is also considered in cellular network technologies (i.e., LTE and 5G) and that is performed through Hybrid Automatic Repeat reQuest (HARQ) techniques. In particular, HARQs imply that instead of retransmitting the entire packet, only a portion of the packet or extra redundant information is retransmitted to demodulate the original packet. In [30], the authors propose the use of a new early HARQ scheme based on LDPC subcodes (SC E-HARQ). This technique decreases the overall latency since enables to start with the feedback calculation before the entire codeword is received. Results indicate that latency values below one millisecond can be obtained for Block Error Rate (BLER) values of less than 10^{-4}. Then, in [31], authors focus their study on delay-constrained scenarios and propose a fast HARQ protocol where in order to decrease the latency some feedback signals and successive messages are disabled. Different transmission rounds are tested and the latency results are reduced up to 60% in the case with five transmissions. Finally, in [32], HARQ techniques are combined with NOMA for the uplink of short packet communications. Simulations include one retransmission per user and coordinated and uncoordinated transmissions. Results show that NOMA-HARQ systems outperform orthogonal multiple access techniques for specific power and latency constraints.

Other authors have also investigated different retransmission schemes, such as [33], where a series of reliability tests using retransmissions are carried out in a laboratory environment. A combination of spatial and frequency diversity over four different RF channels is tested. The results indicate that under certain channel conditions, PER values close to 10^{-7} can be obtained. Later, in [34], different retransmission schemes based on the IEEE 802.15.4 standard are evaluated and compared. The authors conclude that PER values below 10^{-7} are possible with a maximum latency of 3 ms over a one-to-two transmission network with a 30 MHz frequency hop. Finally, in [35], the use of a backup wireless network is proposed to deal with errors. In particular, the authors propose two different models, one based on static redundancy and the other dynamic. The results show that both solutions significantly improve the reliability of the system.

Redundant and spatial diversity based transmissions have led to the development of new communication architectures. One approach is based on "Parallel Redundancy Protocol" (PRP) [36], which was originally developed to achieve seamless redundancy for Ethernet networks requiring high availability. Different works have proposed wireless PRP-like approaches such as [37], where PRP is used as diversity method on the wired Ethernet interfaces of two independent IEEE 802.11 WLAN channels that operate in parallel links. In [38] and [39], the same authors extended their study including the use of PRP for WLAN networks from different points of view such as reliability or latency. Specifically, reliability is measured for different noise levels, different jamming situations, and different packet lengths. In addition, they offer latency and jitter measurements in which they show that the use of PRP considerably improves its performance. Finally, in [40], the use of PRP for sensor networks in a star network configuration is proposed. The MAC level simulations show that when one of the two channels used for redundancy is interfered with, the latency and jitter are improved by 121% and 376%. However, when both channels suffer interference, the improvement is drastically reduced.

Redundancy over Wi-Fi has been recently investigated under the name of Wi-Red (Wi-Fi Redundancy) [41]. Wi-Red aims at providing seamless link-level redundancy in IEEE 802.11 networks for industrial reliable wireless communications. The authors propose two working mechanisms depending on the complexity/reliability tradeoff: Reactive Duplicate Avoidance (RDA) and Proactive Duplicate Avoidance (PDA). From a latency perspective, the authors carried out several tests, where they conclude that both, RDA and PDA, reduce considerably the latency in industrial networks in comparison with other existing solutions. Then, in [42], the authors implemented a Wi-Red prototype using commercial 802.11 devices and they proved with a set of experiments that Wi-Red was able to enhance the latency and reliability in real industrial scenarios.

III. FA USE CASE AND EVALUATION PROCEDURE

This section contains the description of the use case implemented in this work and the evaluation procedure followed to tested the proposed solutions.

A. USE CASE

In this work, the use case emulates a small manufacturing cell with FA requirements [43]. Given that it is a reduced space (i.e., 10 × 10 m), it is assumed that the maximum number of nodes is 20. The network is made up of two types of equipment: Access Points (AP) and nodes. The former is in charge of distributing the information and of the synchronization tasks. On the contrary, the nodes are mainly receivers of the information or the commands sent by the APs. APs and nodes are organized in a centralized topology, where all the nodes receive data from the AP and send back their feedback in a timely organized manner. The diagram of the implemented architecture is shown in Fig. 1, where each node is connected to each AP in the network. Since industrial applications require deterministic communications, the wireless devices are preconfigured to transmit at specific time instants. As depicted in the figure, M indicates the number of nodes, N the number of APs, and γ the period of time between the communication of a node with two different APs. The information transmitted on the network is classified into two types: Critical Service (CS) and Best Effort (BE) service. CS is critical information that requires high reliability rates and low latency. Instead, BE is non-critical information, where the reliability and latency requirements are more flexible. Therefore, CS is configured with the lowest Modulation and Coding Scheme (MCS) (i.e., BPSK 1/2) and BE is configured with higher MCS that provide higher data rates (i.e., QPSK 1/2 - 16QAM 1/2). Regarding the size of the packets, as in [19], the size of the CS information packet has been set at 18 bytes, while for the BE the size is flexible.

Industrial scenarios are characterized by harsh wireless propagation conditions. To perform a realistic evaluation,

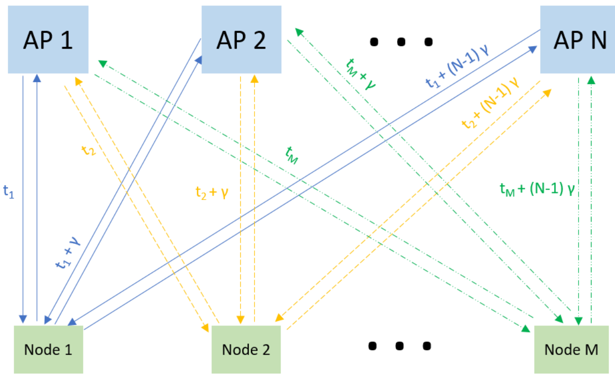


FIGURE 1. Network architecture.

two standard industrial wireless channels have been used, namely, CM7 and CM8 [44], which represents two industrial scenarios with different propagation properties. Although the CM7 and CM8 channels were initially generated based on IEEE 802.15.4a networks, the fact that there are no specific IEEE 802.11n models for industrial environments makes them a suitable candidate for our research purposes. For example, [45] describes several standard channel models for 802.11, but none of them reflects the characteristics of industrial environments. Concerning particular features of CM7 and CM8 channel models, the former represents line-of-sight (LOS) conditions, while CM8 is oriented to non-LOS (NLOS) conditions. As described in [44], channel measurements were carried out in larger enclosures (i.e., factory halls), filled with a large number of metallic reflectors, which implies a severe multipath effect. Moreover, the model was obtained following different measurements that cover a range from 2 to 8 m and frequency band from 2 to 10 GHz. As an example, Fig. 2 shows the channel gain time variability of three uncorrelated realizations of the CM7 channel.

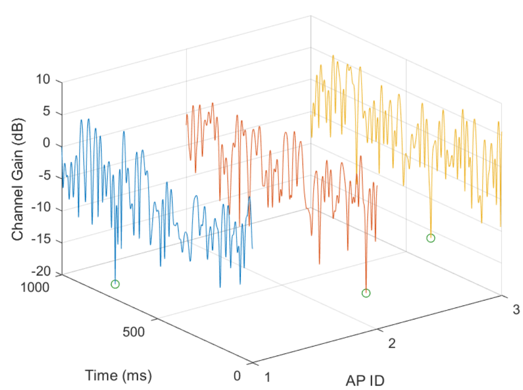


FIGURE 2. CM7 channels obtained from different APs.

The rest of the parameters related to the use case are shown in Table 1.

B. EVALUATION PROCEDURE

The evaluation procedure described in this subsection has been applied to the technical solutions proposed in the

TABLE 1. Use case parameters.

Parameter	Value
Application	FA
Center Frequency	2.4 GHz
Bandwidth	40 MHz
Channel model	CM7 - CM8
Type of Devices	AP and Nodes
Number of Nodes	Up to 20
Cell Size	10 x 10 m
Critical Service (Mbps)	12 Mbps
Non-critical Service (Mbps)	Up to 48 Mbps
Critical Service Payload	18 Byte

following sections. The simulations are based on a software tool designed on purpose for this work. It comprises a mathematical simulator (Matlab) and a network simulator (OMNeT++ [46]). The objective of the first one is the characterization of the reliability under a certain propagation channel while the second one emulates the communication network.

At the beginning of this work, different versions of the 802.11 standard were considered. In the first place, versions older than 802.11n were discarded due to the lack of LDPC codes. Afterwards, the 802.11ac standard was also discarded due to its higher overhead when compared with 802.11n. Finally, the 802.11ax standard presents configurations with similar or even better overhead than 802.11n since it incorporates several novelties, such as the trigger-frame based OFDMA medium access [47]. However, the 802.11ax transceiver presents a much higher complexity than 802.11n and its combination with NOMA may present several challenges. That is why it has been opted to use 802.11n and maintain a tradeoff between complexity and the performance parameters of industrial environments (i.e., reliability and latency) [6].

Concerning the implementation part, in a first step, a transmitter-receiver 802.11n standard chain has been implemented in Matlab [48]. Based on this implementation, several modifications have been carried out in the PHY level to introduce NOMA as a multiplexing option [19]. In addition, the transmitter-receiver chain also implements channel phenomena simulations, including fast-fading, free space loss, etc. To carry out the simulations, packets are sent assuming different channel realizations, and then, in reception, it is evaluated whether the packet has been correctly received or not. Each packet transmission set is performed for different Signal-to-Noise Ratio (SNR) values, with steps of 0.25 dB. The number of simulated packets and the SNR simulation steps have been adapted to the expected PER values. In particular, the simulated packet number is always at least one order of magnitude higher than the required value for obtaining the desired PER value assuming one single error.

In a second step, the network simulation tool is fed with the results obtained with Matlab. In particular, the PHY level performance measurements for different SNR values are introduced in OMNeT++ in order to use them in the error calculation block. Then, the network simulation tool comprises the network level simulation of both PHY and MAC layers. On the one hand, the PHY module facilitates parameters closely related to the physical layer, such as the MCS choice, the length of each transmission slot and the airtime of each data packet. On the other hand, the MAC layer implements the diversity schemes and the medium access control mechanisms. Since OMNeT++ does not support NOMA communications, NOMA has been modelled in OMNeT++ through two independent data flows in the same transmission period. Then, in the reception stage, both data flows are individually and orderly managed (i.e., first data in the UL, then, the data in the LL if UL is successfully recovered).

To decide if a packet is erroneous, this methodology uses the receiver instantaneous SNR. As shown in [49], it is difficult to quantify the uncorrelation grade of two channels with different paths within the same environment, and it is not possible to completely uncorrelate those channels. Therefore, in this work, a case of partial uncorrelation is assumed, where the instantaneous SNR of each node has been modeled as a combination of two components: the mean SNR, which is a static value that depends on the reception characteristics, and the variable attenuation of the channel which varies with time:

$$SNR_{i,k} = \theta_i + \alpha_{i,j,k}, \tag{1}$$

where i identifies the node, j the AP, and k is the time when the transmission takes place. On the other hand, θ is the mean SNR and the variable attenuation of the channel is α. To carry out the simulations with this model, for each path, an independent evolution of the channel variability is assumed (α), where a new channel is generated by varying the generation seed. In Fig. 2, three examples of three different CM7 channels are shown. It should be noted that the three examples have similar behavior since the channel varies with similar rate. However, the minimum fading of each channel (i.e., green circle in the figure) occurs at different moments, which indicates that despite having a similar long-term trend, the instantaneous attenuation at each time point is uncorrelated.

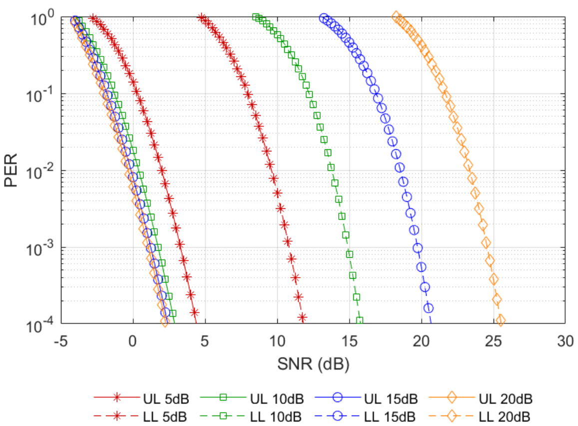
IV. PHY/MAC DESIGN

This section shows, on the one hand, the reliability performance that could be obtained solely with the PHY layer, and, on the other hand, the general architecture of the MAC layer and the retransmission techniques presented afterwards.

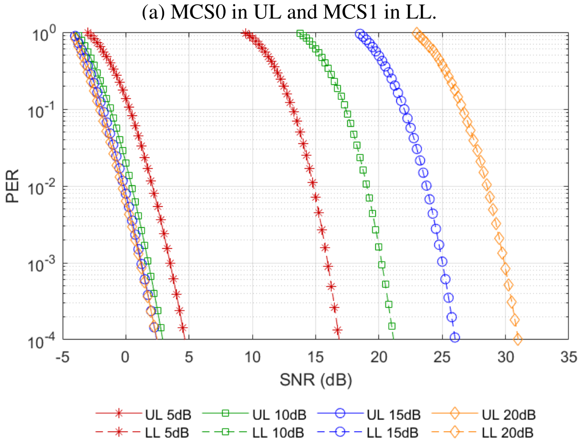
A. PHY LAYER

The development of this section is based on [19], and therefore, the transmitter-receiver chain, as well as the NOMA signal generation and cancellation have not been modified. However, this subsection extends the previous work by focusing on the effect of the IL and the MCS of the Lower Layer (LL). The IL ranges from a case where the LL is injected a few decibels below the Upper Layer (UL) (i.e., 5 dB) to cases where LL is deeply buried (i.e., 20 dB). The LL signal configuration choices are, first, a robust configuration (i.e., MCS1, QPSK 1/2) and, then, a case with doubled capacity and reduced robustness (i.e., MCS3, 16QAM 1/2). On the other hand, the MCS0 is always selected for the UL in order to guarantee high reliability cases. The channel model used in the simulations is the CM7.

Fig. 3, contains the UL performance curves for different injection values (from 5 to 20 dB range). The best result occurs for an IL of 20 dB and provides a SNR requirement of 2.1 dB for a PER value of 10^{-4}. Additionally, UL curves show that the SNR difference in the range [10,20] is insignificant (<1 dB). In consequence, following analyses of the UL behaviour will focus only on the IL range from 5 to 10 dB.



(a) MCS0 in UL and MCS1 in LL.



(b) MCS0 in UL and MCS3 in LL.

FIGURE 3. PHY layer performance for different injection level values and MCS configurations under CM7 propagation channel.

The impact of IL is a two-fold problem because the performance of the LL has to be taken into account as well. The difference among the cases for the LL is more evident. The gap between LL curves is the IL value. Also, for a given IL, the difference between MCS1 and MCS3 is close to 5 dB, which is approximately the performance difference between both MCS configurations. Looking at the LL SNR thresholds for a PER value of 10^{-4}, the results obtained with both 15 and 20 dB lead to SNR thresholds well above 20 dB and look unrealistic for industrial application. IL values of 5 and 10 dB are the most suitable ones since they provide assumable SNR values for the LL.

The same simulations for the CM8 channel model have been carried out and the results are very similar (see Fig. 3). In general, when the CM8 channel model is used, the reliability performance shows a degradation between 0.1 dB and 0.3 dB. Given the similar performance trend in both channel model conditions (CM7 and CM8), the rest of the results presented in the following sections are based only on the CM7 channel model.

B. MAC LAYER

To meet the strict requirements associated with FA environments, PHY techniques have to be combined with efficient MAC level tools. In particular, an adequate MAC layer can potentially improve the reliability, guarantee bounded latency and increase the determinism of the overall communication system. In this section, the description of a MAC layer proposal and the different techniques described in the following sections are introduced.

Fig. 4 shows a time diagram of the superframe structure, which is based on TDMA in order to guarantee deterministic medium access to all the network devices. Three type of blocks can be identified: Downlink CS + BE transmission, Uplink Feedback (UF) and On-demand retransmissions. The first block is the initial transmission of both the critical and best-effort services. This transmission occurs in the downlink, from the AP to each of the nodes. Each downlink slot has a duration of 70 μs because is the minimum time required to transmit the 18 bytes of the CS (see Table 1) by using the lowest and most robust MCS (i.e., MCS0, BPSK 1/2). Furthermore, as the network is made up of 10 nodes the total length of the first block is 700 μs. Then, the uplink feedback phase begins, in which each node informs individually the AP (using each node a dedicated time slot) whether or not the CS and BE services have been correctly received (i.e., ACK/NACK). This information is sent in single-layer mode using the most robust configuration (i.e., MCS0, BPSK 1/2). The duration of each slot is 54 μs since, in this case, only one OFDM symbol is required to transmit the ACK/NACK information. Once the feedback from the nodes is received, the AP retransmits only the packets reported as incorrect. That is, in case the node reports that both services, CS and BE, have been incorrectly received, the AP would retransmit CS and BE in two-layer mode. However, if the node reports that only the BE service is erroneous, the AP retransmits only the BE content in single-layer mode. This block has the same duration as the initial block. Finally, since the number of retransmissions is configurable, blocks two and three would be repeated until the configured number of retransmissions is reached.

Taking into account the superframe structure, the minimum and maximum latency values can be calculated. In the case of the minimum latency, this is obtained when the packet is delivered in the initial transmission block and, therefore, the 70 μs correspond to the duration of a transmission slot.

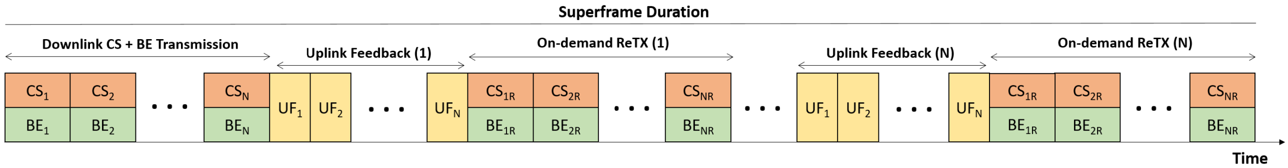


FIGURE 4. Superframe time representation.

On the contrary, the maximum latency value is obtained when the packet is received in the last on-demand retransmissions block. Therefore, the higher number of allowed retransmissions, the higher the maximum latency. That is why, up to three retransmission attempts are evaluated in order to keep assumable maximum latency values. Specifically, the maximum latency values are 1.24 ms, 2.48 ms and 3.72 ms for the one, two and three retransmissions, respectively.

Besides the flexibility in the number of retransmissions, the main difference with the superframe shown in [19] relays on the techniques used to carry out the retransmissions. In this case, in addition to the time-domain (see Section V), retransmissions with spatial diversity will also be implemented (see Section VI), where each one of the retransmissions is transmitted by a different AP. Finally, an additional retransmission mode is proposed. This alternative consists of using a higher injection level in the retransmissions to favor the correct reception of the CS packet (see Section VII). The following sections detail the design of each of the retransmission schemes and evaluate them from the point of view of reliability.

Although the MAC layer configuration presented in this section is specifically designed for the use case introduced in Section III-A, it would be possible to scale the MAC layer configuration to support more complex scenarios. For example, in a more heterogeneous case, there could be nodes that require only the CS, only the BE service, or both. In that case, each node should access the corresponding NOMA layer. If only the CS is needed, once the desired information has been decoded, the node would not access the LL layer. On the contrary, if only the BE service is required, the node should follow the entire decoding process associated with NOMA. However, note that the use of more nodes or with other configurations would not affect the overall reliability performance. Another variation could involve an industrial application built on top of the network architecture that requires a higher uplink traffic. In that case, the superframe could be modified to introduce more specific blocks for the uplink traffic. Furthermore, to introduce more uplink traffic, the medium could be managed through NOMA so that two nodes transmit the information simultaneously. However, this alternative would considerably increase the complexity of the receiver, since the IL between the layers would not be predetermined and would vary in each time slot and, therefore, at the receiver side, a block for estimating the IL would be necessary and/or transmitters should arrange their timing in advance [50].

V. TIME DOMAIN RETRANSMISSIONS

This section contains the design and evaluation of the combination between time-domain retransmissions with the NOMA-based 802.11n communications system.

A. DESIGN

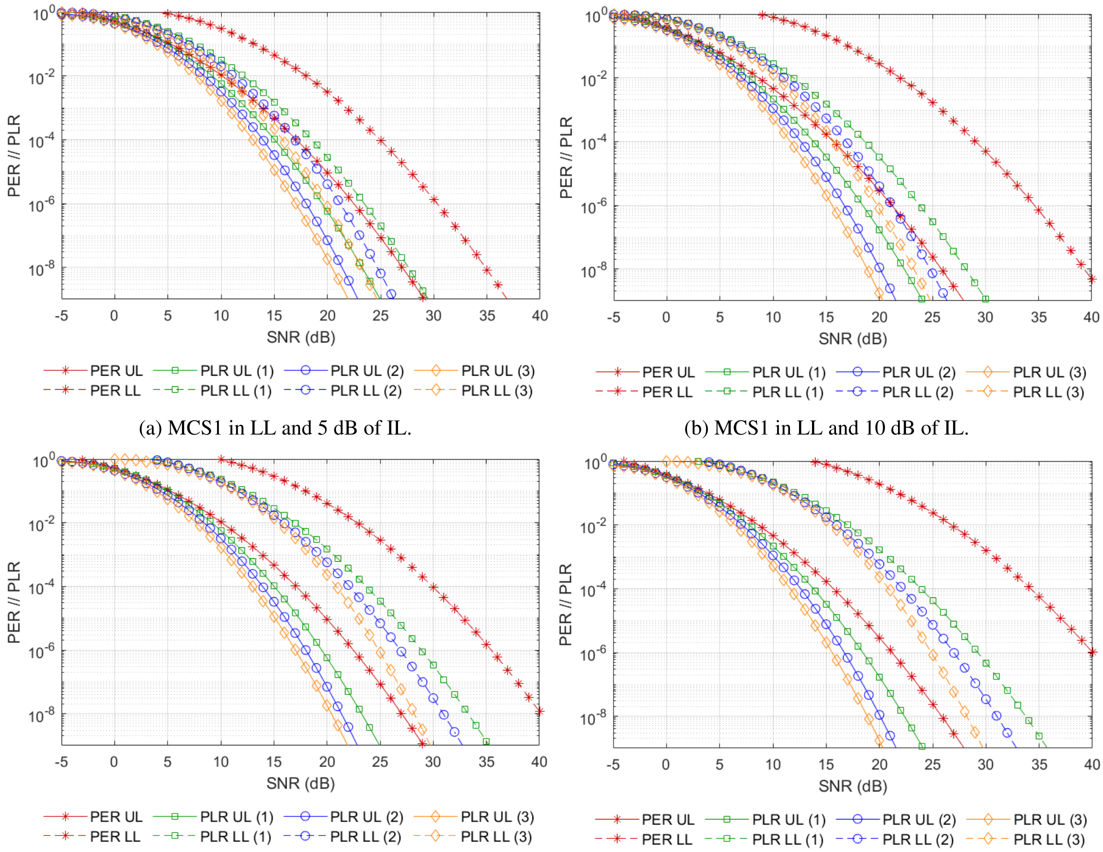
This scheme follows the time diagram presented in Fig. 4, where the blocks of on-demand retransmissions are carried out by the same AP that made the initial attempt (i.e, single transmitter). Consequently, each node SNR for the retransmitted signal is highly correlated with the SNR received in the previous transmission. In particular, following Eq. (1), the only difference between the initial transmission and the retransmission is the temporal evolution of the channel (k), while the origin (j) and the destination (i) is the same.

These types of retransmissions offer better results under channels with high variability and lower time coherence. However, although they are not offering a meaningful gain in all cases in terms of reliability, it should be noted that they entail a very small increase in complexity and implementation cost.

B. EVALUATION

The reliability results obtained with time-domain retransmissions are presented in Fig. 5, which is based on PER and Packet Loss Rate (PLR) measurements. In this case, PER represents the performance when no retransmissions are used, while PLR is for cases with retransmissions. Firstly, the best reliability performance values for the UL are obtained in the 10 dB case, which presents a difference close to 2 dB in comparison with the worst case (i.e., IL 5 dB). When the number of retransmissions is increased, the gap is reduced. In this case, gains close to 4 dB, 6 dB and 7 dB are obtained for one, two and three retransmissions, respectively. Although the reliability performance is improved, the best case SNRs are still around 20 dB, which is considered too high for the critical services. LL shows a different behavior. Although different ILs have been used, the results obtained are very similar, since the retransmissions are done in single-layer mode.

In general, time-domain retransmissions improve reliability but not significantly since the coherence time of the channel is higher than the time between retransmissions. Furthermore, it is worth noting the gain provided by the case of one retransmission, since introducing the first retransmission improves reliability more than when introducing the second or third retransmission. However, it would be necessary to



(a) MCS1 in LL and 5 dB of IL.

(b) MCS1 in LL and 10 dB of IL.

(c) MCS3 in LL and 5 dB of IL.

(d) MCS3 in LL and 10 dB of IL.

FIGURE 5. Reliability results obtained using a single transmitter with time-domain retransmissions under CM7 channel model and MCS0 in UL for different configurations.

evaluate depending on the target application which case best meets the reliability vs latency/complexity relationship.

VI. RETRANSMISSIONS FROM MULTIPLE TRANSMITTERS

This section shows the design and evaluation of combining retransmissions from multiple transmitters with NOMA.

A. DESIGN

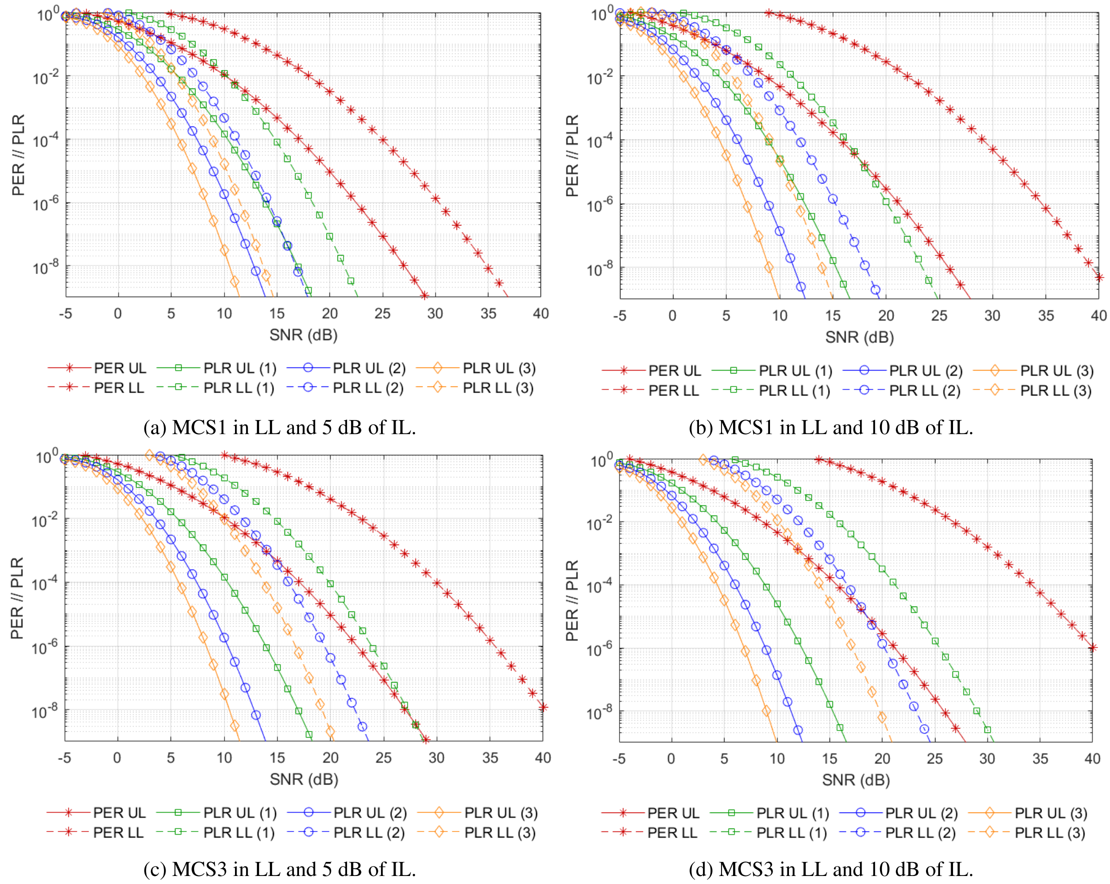
The second technique belongs to the family of spatial diversity and consists of using multiple transmitters. Specifically, taking into account the structure of the superframe (see Fig. 4), each block of on-demand retransmissions will be delivered by a different AP. So if the number of retransmissions configured is N, the number of APs that make up the network has to be $N + 1$. In addition, in the UF period, each node will send its feedback information to the AP that sent it the data in the previous block. It is assumed that all the APs are connected and synchronized so that they all receive the information that the UF contains. Additionally, nodes are suppposed to share a common time reference that grants interference free access to the medium as described in [51], [52].

The retransmissions are conveyed from a different AP, and thus, the initial transmission and the retransmission do not follow the same propagation path. Specifically, according to Eq. (1), since the mean SNR (θ) is affected by the receiver and the implementation environment, it will remain constant for the different retransmissions. On the other hand, when varying the retransmission path using another AP, the characteristics of the variable attenuation will be quite uncorrelated (i.e., j).

Time/space retransmissions from multiple transmitters will improve the reliability offered by time-domain retransmissions. However, introducing this scheme in a network includes higher implementation and synchronization costs. Therefore, depending on the gain obtained in reliability and the type of target application, they could be decisive.

B. EVALUATION

The results are gathered in Fig. 6. First, it is observed that the UL results are considerably better than those obtained using time-domain retransmissions only. In particular, the best results obtained using three retransmissions are around 10 dB. The gain associated to one, two and three retransmission is close to 11, 15 and 18 dB, respectively and for any IL value. The case that implements single retransmission is already better than any of the cases presented in the previous section (see Section V).



(a) MCS1 in LL and 5 dB of IL.

(b) MCS1 in LL and 10 dB of IL.

(c) MCS3 in LL and 5 dB of IL.

(d) MCS3 in LL and 10 dB of IL.

FIGURE 6. Reliability results obtained using retransmissions with multiple transmitters under CM7 channel model and MCS0 in UL for different configurations.

On the other hand, the results obtained for LL present a trend similar to that obtained in the previous section (see Section V). The case with the lowest reliability (i.e., MCS1 and 5 dB of IL) has gains of 14, 19 and 22 dB for one, two and three retransmissions, respectively. The case with the highest gain is that of the MCS3 with an IL of 10 dB, where 18, 24 and 28 dB are achieved for one, two and three retransmissions, respectively.

In general, the reliability obtained using multiple transmitters is greatly improved and the PLR rates obtained are applicable in mission-critical environments. However, an increase in complexity associated with the synchronization and deployment of APs has to be assumed. It is important to note that introducing a second AP is critical, since it provides a considerable increase in reliability. On the contrary, the differential gain provided by a third and fourth AP decreases significantly. Therefore, its implementation will depend on the use case and the reliability that needs to be addressed.

VII. RETRANSMISSIONS WITH VARIABLE IL

This section shows the design and evaluation of the combination of the retransmission techniques shown in the previous sections with adaptive injection levels.

A. DESIGN

This section presents a complementary technique to the previous time/space retransmissions. In this case, an adaptive injection level is proposed to increase the reliability of the UL. The adaptive IL provides a tool to make the UL more robust. If one of the nodes reports in the UF period that both services have been erroneously received, the AP retransmits the packets encoded with a larger IL, so that the LL impact is lower and the SNR required to decode the CS is lower also. However, in case the node has correctly received the CS and only needs the retransmissions of the BE service, the retransmission is carried out in single-layer mode. On the other hand, this technique does not present any additional complexity for the receiver nodes, since as in the UF period each node has requested the corresponding retransmissions if required, each node knows which IL is going to use the AP to encode the information. It should be noted that from now on the acronyms IL1 and IL2 represent the injection level used in the initial transmission and the one used for the retransmissions, respectively.

The adaptive injection level is a complementary technique to those schemes already evaluated in Section V and VI. Although the gain that can be obtained with this technique is limited, the complexity involved is very low, so the reliability/complexity ratio offers an interesting alternative.

B. EVALUATION

Fig. 7 shows the results obtained using the adaptive IL method for all retransmission cases (i.e., time and space) using a different number of retransmissions (i.e., one, two and three). The gain represents the variation of the SNR required to achieve a PLR value of 10^{-9} when introducing the adaptive IL. Given that the gain obtained in all cases is similar and that it does not show a straightforward behavior, all cases have been grouped according to layer and MCS. In Fig. 7 the black whiskers represent the maximum and minimum values of each case, the blue lines represent the 25^{th} and 75^{th} percentiles and the red line the median value. Concerning the IL, for this case, IL1 has been set to 5 dB and IL2 to 10 dB.

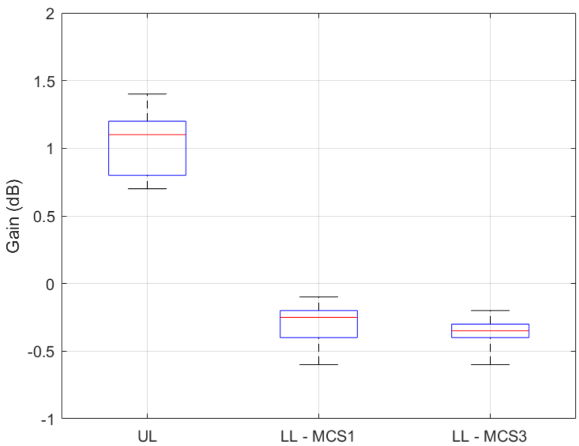


FIGURE 7. Improvement obtained with adaptive IL.

Fig. 7, proves that the performance of the UL layer is improved, assuming small losses in the LL layer. Specifically, the UL shows a median gain of 1.1 dB and taking into account that time retransmissions only provide around 4 dB of improvement with a single retransmission, it is a considerable gain. On the contrary, from the point of view of retransmissions using multiple transmitters, the gain is small compared to the values shown in Section VI-B. On the other hand, the LL reliability degrades. However, taking into account that LL is a BE service, and that degradation is lower than 1 dB, it is considered an acceptable performance loss.

VIII. COMPARISON OF PROPOSED RETRANSMISSION SCHEMES

This section compares all the solutions proposed in this paper in terms of the superframe and network size, reliability, and latency.

A. SUPERFRAME AND NETWORK SIZE

The number of nodes in the network or the duration of the superframe are also parameters to identify different industrial environments. FA, for instance, is more restrictive than Process Automation (PA) in terms of time related parameters such as delay, jitter or update time [43]. Therefore, this section presents the estimation of superframe lengths and the number of nodes in the network using the superframe structure shown in Fig. 4.

First, Table 2 shows the necessary superframe duration for networks in which the number of nodes is pre-established. It should be noted that all cases with a single retransmission have a superframe duration below 10 ms. However, as the number of retransmissions increases, the length of the superframe increases considerably and exceeds FA limits.

TABLE 2. Superframe size for different number of nodes.

Number of Nodes	1 ReTX (ms)	2 ReTX (ms)	3 ReTX (ms)
2	0.39	0.64	0.88
5	0.97	1.59	2.21
15	2.91	4.77	6.63
20	3.88	6.36	8.84
30	5.82	9.54	13.26
40	7.76	12.72	17.68
50	9.70	15.90	22.10

Table 3 displays results from a different approach. We assume that the maximum duration of the superframe is fixed by the use case and limited by the application cycle period. In this case, the table shows the maximum number of nodes that can be included in the superframe. For example, the first case, where the superframe cannot exceed one millisecond, is an FA case where a very small cycle time is required. In that case, the superframe supports up to five, three, and two nodes using one, two, and three retransmissions, respectively. On the other hand, if the limit value of FA environments (i.e., 10 ms) is taken into account, the number of possible nodes is multiplied by ten. Finally, the last two cases are oriented to PA environments and the number of admissible nodes is an order of magnitude higher than in the case of 10 ms.

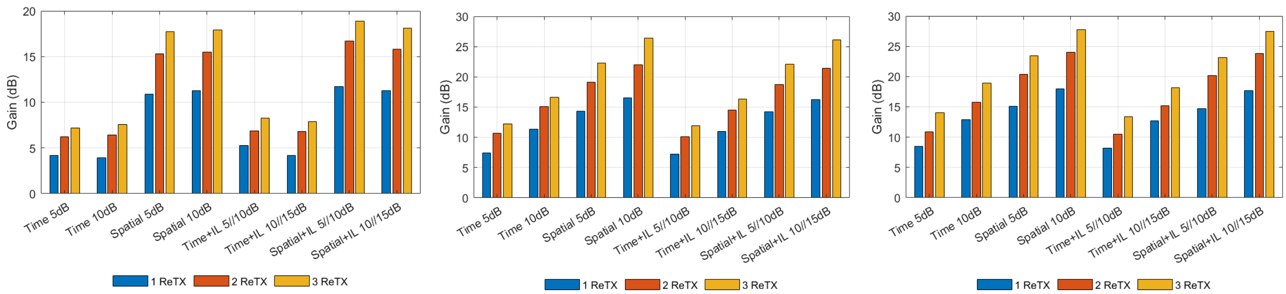
TABLE 3. Maximum number of nodes for different superframe sizes.

Superframe Length	1 ReTX	2 ReTX	3 ReTX
1 ms	5	3	2
5 ms	25	15	11
10 ms	51	31	22
50 ms	257	157	112
100 ms	515	314	226

B. RELIABILITY

In Fig. 8, the gain is calculated as the SNR difference in dB between the PER and the PLR required to have an error rate of 10^{-9}. The analysis is divided into three parts: UL, LL with MCS1 and LL with MCS3, which are represented in Fig. 8(a), Fig. 8(b) and Fig. 8(c), respectively.

In Fig. 8(a), the first conclusion is that retransmissions with multiple transmitters have a greater impact on the UL case. Regarding the adaptive injection level, it is observed that the reliability improves when using it but in a moderate



(a) Gain of the UL. (b) Gain of the LL using MCS1. (c) Gain of the LL using MCS3.

FIGURE 8. Comparison of the gain obtained by the different MAC techniques and for different configurations.

TABLE 4. Mean E2E latency analysis.

SNR	1 ReTX (μs)				2 ReTX (μs)				3 ReTX (μs)			
	Time	Sp.	Time + IL	Sp. + IL	Time	Sp.	Time + IL	Sp. + IL	Time	Sp.	Time + IL	Sp. + IL
UL												
5 dB	118	207	157	216	172	241	200	234	235	248	245	237
10 dB	76	83	80	83	83	83	84	83	88	83	86	83
15 dB	70	70	70	70	71	70	70	70	71	70	71	70
LL using MCS1												
5 dB	960	950	962	977	1048	1281	1045	1218	1145	1424	1139	1311
10 dB	451	466	453	468	476	492	475	488	498	494	496	489
15 dB	132	133	132	133	134	133	134	133	135	133	135	133
LL using MCS3												
5 dB	1226	1224	1226	1226	1406	1800	1407	1751	1605	2293	1605	2213
10 dB	970	983	970	985	1057	1242	1057	1237	1155	1348	1155	1339
15 dB	461	478	462	478	485	500	486	500	507	501	508	501

way. Finally, increasing the number of retransmissions is more effective over retransmissions with spatial diversity, with overall gain values close to 20 dB in some cases.

On the other hand, in Fig. 8(b) and Fig. 8(c), the LL shows higher gains than the UL, due to the configuration of retransmissions in single-layer mode. The performance behavior depending on the configuration is quite similar, although in general, the LL configured with MCS3 shows a slightly higher gain than MCS1. Furthermore, in this case, although the difference in performance between time-domain retransmissions and with multiple transmitters is still evident, it has diminished considerably. On the other hand, it can be said that the use of the adaptive injection level does not affect the LL, since the losses are negligible. Finally, it is worth noting that there are several cases in both MCS1 and MCS3 that exceed 25 dB of gain.

C. E2E LATENCY

Latency is another critical parameter in FA environments, which in general is related to reliability since many of the techniques to improve reliability involve an increase in latency. Therefore, in this work latency and reliability are analyzed to give a global vision of performance. In this case, End-to-End (E2E) latency is used to represent the time that elapses since the transmission of a packet begins until it is finally received, including retransmissions if necessary. The most representative E2E latency values have been gathered in Table 4. It shows the mean E2E latency values for different SNR values and each of the retransmission techniques. Specifically, the values shown for retransmissions in the time domain and with multiple transmitters have been obtained using an IL of 5 dB and those that use the adaptive injection level vary between 5 dB and 10 dB. Those cases have been chosen because they present the most balanced reliability results between UL and LL in the previous subsection. To calculate the mean latencies, three specific SNR values have been selected (i.e., 5, 10 and 15 dB). On average, the UL presents low latency values in which 250 μs are not exceeded in any case. In fact, the latency values with 15 dB SNR are practically the same as the minimum value, which indicates a high reliability rate. In turn, as expected, the latency values for LL are higher than for UL. However, despite the low SNR values used, in no case do they exceed 2 ms and taking into account that it is a BE service, it is an acceptable rate. Furthermore, it should be noted that for the 10 dB and 15 dB SNR cases using MCS1 and the 15 dB case using MCS3, the maximum mean latency values are around 500 μs.

Finally, if the latency values are analyzed as a function of the retransmission technique, it can be seen that the higher the reliability rate, the higher the average latency that must be assumed. That is why, the highest values appear in the case of multiple transmitters combined with adaptive IL since it is the case in which retransmissions are more efficient. The main reason for this is the success rate of the retransmission

TABLE 5. Analysis of the jitter results obtained for the UL.

SNR	1 ReTX (μs)				2 ReTX (μs)				3 ReTX (μs)			
	Time	Sp.	Time + IL	Sp. + IL	Time	Sp.	Time + IL	Sp. + IL	Time	Sp.	Time + IL	Sp. + IL
5 dB	114.37	114.06	114.43	114.26	194.57	168.27	194.16	168.41	226.62	203.50	226.21	203.54
10 dB	16.24	17.84	16.19	17.88	24.59	21.58	24.41	21.51	30.81	23.49	30.89	23.54
15 dB	0.81	0.82	0.77	0.85	0.98	0.88	0.96	0.89	1.12	0.87	1.06	0.89

schemes since error packets are not taken into account for the latency calculation. However, taking into account the gains in reliability that have been obtained in the previous section, the increase in latency is acceptable. Similarly, the more retransmissions that are used, the greater the latency that is obtained, since the size of the superframe is also lengthened.

To complement the E2E latency measurements, Table 5 shows the jitter values obtained for the UL with the same configurations as in Table 4. First of all, it should be noted that with an SNR of 5 dB the jitter values are quite high and even in some cases very close to the average latency. However, the jitter is greatly reduced when the SNR is increased. In particular, with 10 dB of SNR the jitter does not exceed 30 μs and with 15 dB the values are around one μs, which indicates that hardly any retransmissions are needed. Furthermore, in general, retransmissions based on multiple transmitters have better jitter values, since as they are more efficient, they require fewer retransmissions to guarantee the correct reception of the packets.

In summary, depending on the requirements of the final application, a compromise between latency and reliability has to be made.

IX. CONCLUSION

This paper presents and evaluates a set of techniques to increase communication reliability over the 802.11n standard. These techniques include the integration of NOMA within the 802.11n PHY layer, the use of a TDMA-based MAC scheme, and the use of different retransmission techniques in different diversity domains. To our best knowledge, this paper is the first to consider different retransmission schemes for a NOMA-based 802.11n communication system. In particular, time-domain and multiple transmitter based retransmissions have been combined with adaptive injection level and evaluated from the reliability and latency point of view.

The main conclusion is that each one of the evaluated retransmission techniques has its advantages and disadvantages. In particular, time-domain retransmissions are the simplest to implement, but also the ones that offer lower reliability due to channel dependence. However, it has been shown that by increasing the number of retransmissions, up to 7 dB of gain can be obtained for the CS. In contrast, retransmission schemes that include multiple transmitters have demonstrated much higher reliability rates. The main disadvantages of these techniques are the synchronization that is required and the deployment costs. The specific NOMA technique aimed at adapting the injection level to favor the reception of the UL in retransmissions has shown moderate gain values, but it can be useful in cases where the highest reliability rate is needed. Moreover, the latency analysis indicates that the higher the reliability offered by the retransmission schemes, the higher the latency that the system has to assume. However, taking into account the type of services proposed and the requirements of industrial environments, the latency values obtained are below the limits. Therefore, depending on the final application, in order to decide the optimal solution, a tradeoff has to be evaluated and assumed between reliability and latency.

This paper is oriented to the combination of NOMA techniques with TDMA-based medium access, therefore, the next step is to integrate NOMA in OFDMA-based communication schemes such as 802.11ax/be standards. To do this, on the one hand, the necessary modifications in the current 802.11ax/be transceiver architecture will be studied, as well as the potential gains that could be obtained. In addition, the complexity implications of the introduction of NOMA should be measured and low-complexity alternatives should be proposed to improve the complexity/performance tradeoff.

REFERENCES

[1] R. Drath and A. Horch, "Industrie 4.0: Hit or hype? [industry forum]," *IEEE Ind. Electron. Mag.*, vol. 8, no. 2, pp. 56–58, Jun. 2014.

[2] A. Varghese and D. Tandur, "Wireless requirements and challenges in industry 4.0," in *Proc. Int. Conf. Contemp. Comput. Informat. (IC I)*, Nov. 2014, pp. 634–638.

[3] P. Schulz, M. Matthe, H. Klessig, M. Simsek, G. Fettweis, J. Ansari, S. A. Ashraf, B. Almeroth, J. Voigt, I. Riedel, A. Puschmann, A. Mitschele-Thiel, M. Müller, T. Elste, and M. Windisch, "Latency critical IoT applications in 5G: Perspective on the design of radio interface and network architecture," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 70–78, Feb. 2017.

[4] K. Montgomery, R. Candell, Y. Liu, and M. Hany, *Wireless User Requirements for the Factory Workcell* (Advanced Manufacturing Series (NIST AMS)), vol. 8. Gaithersburg, MD, USA: NIST, 2019.

[5] O. Seijo, I. Val, and J. A. Lopez-Fernandez, "w-SHARP: Implementation of a high-performance wireless time-sensitive network for low latency and ultra-low cycle time industrial applications," *IEEE Trans. Ind. Informat.*, early access, Jul. 7, 2020, doi: 10.1109/TII.2020.3007323.

[6] F. Tramarin, S. Vitturi, M. Luvisotto, and A. Zanella, "On the use of IEEE 802.11n for industrial communications," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1877–1886, Oct. 2016.

[7] R. Rondón, M. Gidlund, and K. Landernäs, "Evaluating Bluetooth low energy suitability for time-critical industrial IoT applications," *Int. J. Wireless Inf. Netw.*, vol. 24, no. 3, pp. 278–290, Sep. 2017.

[8] M. Anwar, Y. Xia, and Y. Zhan, "TDMA-based IEEE 802.15.4 for low-latency deterministic control applications," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 338–347, Feb. 2016.

[9] I. Aktas, M. H. Jafari, J. Ansari, T. Dudda, S. A. Ashraf, and J. C. S. Arenas, "LTE evolution—Latency reduction and reliability enhancements for wireless industrial automation," in *Proc. IEEE 28th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Oct. 2017, pp. 1–7.

[10] X. Jiang, M. Luvisotto, Z. Pang, and C. Fischione, "Latency performance of 5G new radio for critical industrial control systems," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2019, pp. 1135–1142.

[11] G. Hampel, C. Li, and J. Li, "5G ultra-reliable low-latency communications in factory automation leveraging licensed and unlicensed bands," *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 117–123, May 2019.

[12] D. Lopez-Perez, A. Garcia-Rodriguez, L. Galati-Giordano, M. Kasslin, and K. Doppler, "IEEE 802.11be extremely high throughput: The next generation of Wi-Fi technology beyond 802.11ax," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 113–119, Sep. 2019.

[13] A. Slalmi, R. Saadane, A. Chehri, and H. Kharraz, "How will 5G transform industrial IoT: Latency and reliability analysis," in *Human Centred Intelligent Systems*. Singapore: Springer, 2020, pp. 335–345.

[14] M. Luvisotto, Z. Pang, and D. Dzung, "Ultra high performance wireless control for critical applications: Challenges and directions," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1448–1459, Jun. 2017.

[15] M. Luvisotto, Z. Pang, and D. Dzung, "High-performance wireless networks for industrial control applications: New targets and feasibility," *Proc. IEEE*, vol. 107, no. 6, pp. 1074–1093, Jun. 2019.

[16] O. Seijo, Z. Fernandez, I. Val, and J. A. Lopez-Fernandez, "SHARP: A novel hybrid architecture for industrial wireless sensor and actuator networks," in *Proc. 14th IEEE Int. Workshop Factory Commun. Syst. (WFCS)*, Jun. 2018, pp. 1–10.

[17] E. Iradier, J. Montalban, L. Fanari, and P. Angueira, "On the use of spatial diversity under highly challenging channels for ultra reliable communications," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2019, pp. 200–207.

[18] M. C. Lucas-Estan, B. Coll-Perales, and J. Gozalvez, "Redundancy and diversity in wireless networks to support mobile industrial applications in industry 4.0," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 311–320, Jan. 2021.

[19] J. Montalban, E. Iradier, P. Angueira, O. Seijo, and I. Val, "NOMA-based 802.11n for industrial automation," *IEEE Access*, vol. 8, pp. 168546–168557, 2020.

[20] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.

[21] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1616–1626, Apr. 2008.

[22] F. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative detectors for trellis-code multiple-access," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1478–1485, Sep. 2002.

[23] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. Elkashlan, I. Chih-Lin, and H. V. Poor, "Application of non-orthogonal multiple access in LTE and 5G networks," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.

[24] E. Iradier, J. Montalban, L. Fanari, P. Angueira, L. Zhang, Y. Wu, and W. Li, "Using NOMA for enabling broadcast/unicast convergence in 5G networks," *IEEE Trans. Broadcast.*, vol. 66, no. 2, pp. 503–514, Jun. 2020.

[25] N. M. Balasubramanya, A. Gupta, and M. Sellathurai, "Combining code-domain and power-domain NOMA for supporting higher number of users," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[26] L. Zhang, W. Li, Y. Wu, X. Wang, S.-I. Park, H. M. Kim, J.-Y. Lee, P. Angueira, and J. Montalban, "Layered-division-multiplexing: Theory and practice," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 216–232, Mar. 2016.

[27] E. Iradier, J. Montalban, L. Fanari, P. Angueira, O. Seijo, and I. Val, "NOMA-based 802.11n for broadcasting multimedia content in factory automation environments," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Jun. 2019, pp. 1–6.

[28] M. Jonsson and K. Kunert, "Towards reliable wireless industrial communication with real-time guarantees," *IEEE Trans. Ind. Informat.*, vol. 5, no. 4, pp. 429–442, Nov. 2009.

[29] P. G. Peon, E. Uhlemann, W. Steiner, and M. Bjorkman, "Applying time diversity for improved reliability in a real-time heterogeneous MAC protocol," in *Proc. IEEE 85th Veh. Technol. Conf. (VTC Spring)*, Jun. 2017, pp. 1–7.

[30] B. Goektepe, S. Faehse, L. Thiele, T. Schierl, and C. Hellge, "Subcode-based early HARQ for 5G," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.

[31] B. Makki, T. Svensson, G. Caire, and M. Zorzi, "Fast HARQ over finite blocklength codes: A technique for low-latency reliable communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 194–209, Jan. 2019.

[32] 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.*, early access, Oct. 2, 2020, doi: 10.1109/JIOT.2020.3028434.

[33] H. Beikirch, M. Voss, and A. Fink, "Redundancy approach to increase the availability and reliability of radio communication in industrial automation," in *Proc. IEEE Conf. Emerg. Technol. Factory Autom.*, Sep. 2009, pp. 1–4.

[34] B. Kilberg, C. B. Schindler, A. Sundararajan, A. Yang, and K. S. J. Pister, "Experimental evaluation of low-latency diversity modes in IEEE 802.15.4 networks," in *Proc. IEEE 23rd Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, vol. 1, Sep. 2018, pp. 211–218.

[35] P. G. Peon, W. Steiner, and E. Uhlemann, "Network fault tolerance by means of diverse physical layers," in *Proc. 25th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, vol. 1, Sep. 2020, pp. 1697–1704.

[36] H. Kirrmann, M. Hansson, and P. Muri, "IEC 62439 PRP: Bumpless recovery for highly available, hard real-time industrial networks," in *Proc. IEEE Conf. Emerg. Technol. Factory Autom. (EFTA)*, Sep. 2007, pp. 1396–1399.

[37] M. Rentschler and P. Laukemann, "Towards a reliable parallel redundant WLAN black channel," in *Proc. 9th IEEE Int. Workshop Factory Commun. Syst.*, May 2012, pp. 255–264.

[38] M. Rentschler, O. A. Mady, M. T. Kassis, H. H. Halawa, T. K. Refaat, R. M. Daoud, H. H. Amer, and H. M. ElSayed, "Simulation of parallel redundant WLAN with OPNET," in *Proc. IEEE 18th Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2013, pp. 1–8.

[39] M. Rentschler and P. Laukemann, "Performance analysis of parallel redundant WLAN," in *Proc. IEEE 17th Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2012, pp. 1–8.

[40] M. Hendawy, M. ElMansoury, K. N. Tawfik, M. M. ElShenawy, A. H. Nagui, A. T. ElSayed, H. H. Halawa, R. M. Daoud, H. H. Amer, M. Rentschler, and H. M. ElSayed, "Application of parallel redundancy in a Wi-Fi-based WNCS using OPNET," in *Proc. IEEE 27th Can. Conf. Electr. Comput. Eng. (CCECE)*, May 2014, pp. 1–6.

[41] G. Cena, S. Scanzio, and A. Valenzano, "Seamless link-level redundancy to improve reliability of industrial Wi-Fi networks," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 608–620, Apr. 2016.

[42] G. Cena, S. Scanzio, and A. Valenzano, "A prototype implementation of Wi-Fi seamless redundancy with reactive duplication avoidance," in *Proc. IEEE 23rd Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2018, pp. 179–186.

[43] J. F. Coll, J. Chilo, and B. Slimane, "Radio-frequency electromagnetic characterization in factory infrastructures," *IEEE Trans. Electromagn. Compat.*, vol. 54, no. 3, pp. 708–711, Jun. 2012.

[44] A. F. Molisch, K. Balakrishnan, C.-C. Chong, S. Emami, A. Fort, J. Karedal, J. Kunisch, H. Schantz, U. Schuster, and K. Siwiak, "IEEE 802.15. 4a channel model-final report," *IEEE P802*, vol. 15, no. 4, p. 0662, Nov. 2004.

[45] V. Erceg, L. Schumacher, and P. Kyritsi, *TGN Channel Models*, document IEEE 802.11-03/940r1, Garden Grove, CA, USA, 2004.

[46] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proc. 1st Int. ICST Conf. Simulation Tools Techn. Commun. Netw. Syst.*, 2008, p. 60.

[47] B. Bellalta, "IEEE 802.11ax: High-efficiency WLANS," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 38–46, Feb. 2016.

[48] *Orthogonal Frequency Division Multiplexing (OFDM) PHY Specification, de Standard—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Nueva York, Estados Unidos*, Standard 802.11, IEEE Computer Society, 2012, pp. 1583–1630.

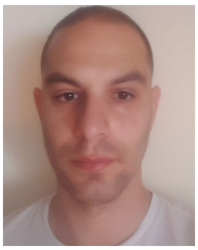
[49] V. N. Swamy, P. Rigge, G. Ranade, B. Nikolic, and A. Sahai, "Wireless channel dynamics and robustness for ultra-reliable low-latency communications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 705–720, Apr. 2019.

[50] S. Abeywickrama, L. Liu, Y. Chi, and C. Yuen, "Over-the-air implementation of uplink NOMA," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

[51] O. Seijo, J. A. Lopez-Fernandez, H.-P. Bernhard, and I. Val, "Enhanced timestamping method for subnanosecond time synchronization in IEEE 802.11 over WLAN standard conditions," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5792–5805, Sep. 2020.

[52] A. Mahmood, R. Exel, H. Trsek, and T. Sauter, "Clock synchronization over IEEE 802.11—A survey of methodologies and protocols," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 907–922, Apr. 2017.

ENEKO IRADIER (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunications engineering from the University of the Basque Country (UPV/EHU), in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree with the TSR Research Group. Since 2015, he is part of the TSR Research Group, UPV/EHU. He was with the Communications Systems Group, IK4-Ikerlan, as a Researcher, from 2017 to 2018. During his doctoral studies, he did an internship with the Communications Research Centre Canada, Ottawa. His current research interests include design and development of new technologies for the physical layer of communication systems and wireless solutions for Industry 4.0. He has served as a reviewer for several renowned international journals and conferences in the area of wireless communications.



LORENZO FANARI (Graduate Student Member, IEEE) received the B.Sc. degree in electrical and electronic engineering and the M.Sc. degree in telecommunication engineering from the University of Cagliari, Italy, in 2015 and 2018, respectively. He is currently pursuing the Ph.D. degree with the University of Basque Country, Spain. His research interests include coding theory and wireless communications.

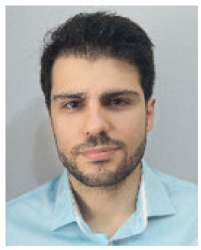
IÑIGO BILBAO (Graduate Student Member, IEEE) received the B.S. degree in physics and the M.S. in space science and technology from the University of the Basque Country (UPV/EHU), in 2019 and 2020, respectively, where he is currently pursuing the Ph.D. degree with the TSR Research Group. Since 2020, he is part of the TSR Research Group, UPV/EHU. His current research interests include design and development of new technologies for the physical layer of wireless solutions for Industry 4.0.

JON MONTALBAN (Senior Member, IEEE) received the M.S. and Ph.D. degrees in telecommunications engineering from the University of the Basque Country, Spain, in 2009 and 2014, respectively. He is part of the TSR (Radiocommunications and Signal Processing) Research Group with the University of the Basque Country, where he is an currently an Assistant Professor involved in several research projects. He has hold visiting research appointments with the Communication Research Centre (CRC), Canada, and Dublin City University (DCU), Ireland. His current research interests include wireless communications and signal processing for reliable industrial communications. He was a co-recipient of several best paper awards, including the Scott Helt Memorial Award to recognize the best paper published in the IEEE Transactions on Broadcasting, in 2019. He has served as a reviewer for several renowned international journals and conferences in the area of wireless communications. He is currently serves as an Associate Editor for IEEE Access.

PABLO ANGUEIRA (Senior Member, IEEE) received the M.S. and Ph.D. degrees in telecommunication engineering from the University of the Basque Country, Spain, in 1997 and 2002, respectively.

In 1998, he joined the Department of Communications Engineering, University of the Basque Country, where he is currently a Full Professor. He is part of the staff of the Signal Processing and Radiocommunication Laboratory, where he has been involved in research on digital broadcasting (DVB-T, DRM, T-DAB, DVB-T2, DVB-NGH, and ATSC 3.0) for more than 20 years. He is a coauthor of an extensive list of articles in international peer-reviewed journals, and a large number of conference presentations in digital broadcasting. He has also coauthored several contributions to the ITU-R working groups WP6 and WP3. His main research interest includes design and development of new technologies for the physical layer of communication systems in industrial wireless environments.

Dr. Angueira is an Associate Editor of the IEEE Transactions on Broadcasting, a member of the IEEE BMSB International Steering Committee, and a Distinguished Lecturer of the IEEE BTS. He serves on the Administrative Committee for the IEEE BTS.



OSCAR SEIJO (Member, IEEE) received the B.Sc. and M.Sc. degrees in telecommunications engineering from the University of Oviedo, Spain, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Communications Systems Group, IKERLAN Technological Research Centre, Mondragón, Spain, in collaboration with the Signal Theory and Communications (TSC) Group, University of Oviedo. His research interests include wireless high-performance PHY and MAC design for industrial applications, time synchronization over wireless systems, and digital signal processing.

IÑAKI VAL (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees from the Department of Electronics Engineering, University of Mondragon, Spain, in 1998 and 2001, respectively, and the Ph.D. degree from the Department of Signals, Systems, and Radiocommunication, Polytechnic University of Madrid, Spain, in 2011. Since 2001, he has been with the Communications Systems Group, IKERLAN, Mondragón, Spain. He was with the Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany, as an Invited Researcher, from 2005 to 2006. He is currently the Team Leader of Communication Systems Group. His research interests include design and implementation of digital wireless communications systems, industrial real-time requirements, communications for distributed control systems, vehicular communications, time synchronization, wireless channel characterization, and digital signal processing. He is currently focused on industrial communication applications.