MUSA Grant-Free Access Framework and Blind Detection Receiver

Guilherme Aquino, Tiago Barbosa, Marwa Chafii, Luciano Mendes and Abdul Karim Gizzini

Abstract—Recently, a non-orthogonal multiple access scheme called multi-user shared access (MUSA) was proposed to provide massive connection capability of low-complexity devices in the 5G networks. MUSA achieves higher spectral efficiency allowing independent devices to transmit data on the same physical layer time-frequency resources. Furthermore, MUSA introduces a grant-free transmission and a blind multi-user detection at the receiver, reducing the complexity on the transmit side. This approach is interesting for Internet of Things applications over mobile communication networks, where the devices have limited power and processing capacity. The references available in the literature about this multiple access scheme do not bring sufficient details about the MUSA multi-user detector. This limitation makes it difficult to evaluate the MUSA performance and to propose improvements for this new technique. The main goal of this paper is to provide a framework describing the entire communication chain using MUSA as multiple access. This paper also brings a proposal for a blind multi-user detection, where the information about the MUSA parameters and the channel state information are unknown at the receiver side. The performance of the MUSA multi-user detector is improved by a deep learning based processing that enhances the quality of the channel estimation provided by a initial minimum mean square error estimator. The proposed deep neural network architecture employed to improve the channel estimation allows more users to share the same time-frequency resources for a given target block error rate, increasing the overall spectrum efficiency of the system.

Index Terms—Grant-free access, blind multi-user detection, MUSA, NOMA, successively interference cancellation, mMTC, deep neural network.

I. INTRODUCTION

THE Fifth Generation of Mobile Network (5G) has four different application scenarios, each one with specific key requirements. The enhanced Mobile Broadband (eMBB) [1] provides high throughput connectivity, aiming for 10 times the data provided by Long-Term Evolution (LTE) networks. The Ultra Reliable Low Latency Communications (URLLC) [2] supports robust low latency connectivity with a target Round Trip Time (RTT) of 1 ms. The enhanced Remote Area Communications (eRAC) [3] provides connectivity in long-range and remote areas, achieving at least 100 Mbps at 50 km from the Base Station (BS) while exploiting the TV White Space (TVWS). Finally, the massive Machine Type Communications (mMTC) [4] promises connectivity to a plethora of power-limited Internet of Things (IoT) devices.

Although all applications scenarios imposes challenging requirements for the 5G network, accommodating a large number of IoT devices, which are unable to waste power on pilots signalling, on a limited number of Physical Layer (PHY) time-frequency resources deserves special attention. The main challenge for the mMTC scenario is to allow a massive number of power-limited and complexity-restricted devices to connect to the network. These IoT devices are typically sensors that transmit low amount of data and cannot afford the cumbersome synchronization process or the pilot signaling required by Orthogonal Multiple Access (OMA). The data traffic generated by these devices has a specific profile that can be exploited by innovative multiple access technique. Usually, the traffic generated by IoT devices is quite different from the human-driven communications scenarios. Typically, in mMTC applications, there is no severe restriction in the latency, the required data rate is very low and the number of transmission per device per day is small [5]. But a high number of devices are expected to be connected to the network, which means that the 5G BS must be able to handle several simultaneous transmissions from different limited devices [6].

Conventional OMA schemes will require large bandwidth to accommodate all transmissions, contradicting the premise of spectral efficiency for 5G systems [7]. Different Non-Orthogonal Multiple Access (NOMA) schemes have been considered as a potential solution for multiple access in IoT applications over 5G networks. NOMA techniques allows for massive connectivity without bandwidth increment due to the overloading concept [8]. In the overloaded Radio Access Network (RAN) [9], the data symbol of different users can be superimposed within the time-frequency resources, increasing the capacity and spectral efficiency of the system. NOMA schemes can control the mutual interference among the users at the cost of complexity increment at the BS [10]. There are two categories for the NOMA schemes in the literature: the power-domain NOMA [11] and the code-domain NOMA [12]. Among the code-domain NOMA the Sparse Code Multiple Access (SCMA) [13], Pattern Division Multiple Access (PDMA) [14], Interleaved Division Multiple Access (IDMA) [15], Lattice Partition Multiple Access (LPMA) [16] and Multi-User Shared Access (MUSA) [17] are the most

G. Aquino, T. Barbosa and L. Mendes are with the Instituto Nacional de Telecomunicações - Inatel, Santa Rita do Sapucaí, MG, Brazil. E-mails: guilhermeaquino@inatel.br, tiago.cardoso@inatel.br and lucianol@inatel.br.

Marwa Chafii and Abdul Karim Gizzini are with ETIS, UMR8051, CY Cergy Paris Université, ENSEA, CNRS, France. E-mails: marwa.chafii@ensea.fr and abdulkarim.gizzini@ensea.fr

promising solutions.

For IoT applications, MUSA becomes an attractive solution because it is a grant-free multiple access conceived for the mMTC scenario [17]. In this multiple access scheme, the data of a given user is spread with a short length code that belongs to a family of the non-orthogonal complex spreading sequences. The data symbol from different users are superposed in a set of time-frequency resources. As a grant-free access scheme, each user can choose its spreading code autonomously, eliminating the signaling for access schedule. Furthermore, no pilots or training sequences are used in the frame structure, resulting in a higher spectral efficiency. This approach means that the receiver has to blindly detect the users' data symbols, without the explicit Channel State Information (CSI). A blind detection based on Successive Interference Cancellation (SIC) can be used to implement the Multi-user Detector (MUD). Since no pilot signaling are transmitted, MUSA can reduce the overhead and the energy consumption of IoT devices, while providing high spectrum efficiency at the cost of higher complexity at the BS.

MUSA is receiving attention from the research community and several contributions have been presented after the introduction of this multiple access scheme, where the structures for transmitter and receiver have been presented, as well as the design of the complex spreading code [17]. Numerical results have shown that MUSA can achieve acceptable Block Error Rate (BLER) performance in very high overloading conditions [18]. However, in [17] and [18], an ideal receiver is used for the performance evaluation. It is assumed that the BS knows the CSI and the spreading sequences used by the IoT devices. These premises are incompatible with a grant-free multiple access with blind detection. Under this unrealistic assumption, the data of different users can be successfully decoded even in highly overloaded situations. This same unrealistic assumption is used by [10] for comparisons between MUSA and other NOMA schemes. In [12] and [19], the authors superficially describes the MUSA scheme without presenting implementation details. In [20], a more realistic receiver implementation was presented and the MUSA BLER performance was analyzed assuming flat fading channel and blind detection receiver. In [21], the BLER performance of a realistic blind detection receiver for grant-free MUSA is analyzed, assuming a frequency selective channel where no pilot nor preamble are used for channel estimation. The numerical results have shown that realistic grant-free MUSA with blind detection achieves an acceptable BLER performance even for a very large number of users colliding in the same transmission resource. Finally, [22] brings a performance analysis similar to the ones presented in [20] and [21]. However, [20], [21] and [22] do not describe how the receivers were implemented and how the SIC removes the interference. In other words, [20], [21] and [22] do not show how the receiver can decode the data of a given device without knowing the CSI and the spreading sequence used in transmission. Without these details, the results cannot be properly reproduced and improvements for MUSA cannot be proposed.

The aim of this paper is to provide a complete framework of how to implement a grant-free MUSA scheme with details

of the transmitter and receiver with blind MUD based on SIC. The channel estimation on the receiver side will be improved by an Artificial Intelligence (AI) algorithm based on deep neural network (DNN) [23]. The DNN is previously trained with a set of Least Square (LS) estimated channel gains at high Signal-to-Noise Ratio (SNR) and low overloading factor. The trained DNN is then used to improve the channel estimation for a range of overloading factors.

The remaining of this paper is organized as follows: Section II presents the MUSA transmitter, while Section III shows the details of the blind MUD based on SIC. This section also describes how the DNN has been used to improve the channel estimation. Section IV brings the MUSA performance evaluation based on numerical simulation for different overload values. Finally, Section V concludes this paper.

II. Principles of the MUSA transmitter

MUSA can accommodate a large number of devices in a limited number of time-frequency resources. The ratio between the number of devices and the number of time-frequency resources in the PHY layer is defined as overloading factor. In order to achieve high overloading factors, the data symbol of a given device is spread over a set of time-frequency resources by a complex sequence from a codebook, which is known at the receiver side. Other devices spread their data symbols using spread sequences from this codebook and transmit it using the same time-frequency resources. This procedure is similar to Direct Sequence Spread Spectrum (DS-SS) employed in Code Division Multiple Access (CDMA) [24]. However, instead of using orthogonal binary codes, MUSA employs non-orthogonal short length complex spread sequences. The authors in [17] have proposed a tri-level complex spreading code with real and imaginary parts in the set $\{-1, 0, 1\}$, leading to the Cartesian representation depicted in Fig. 1.
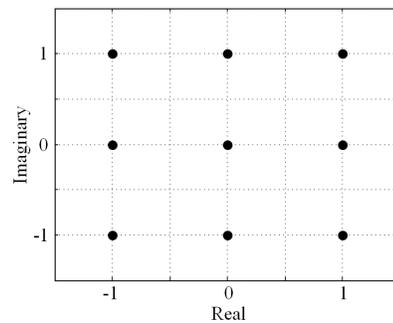
Fig. 1. Cartesian representation of the components of the tri-level complex spreading sequences.

The number of different codes that can be obtained from this alphabet is 9^L, where L is the length of the codeword. From this total of spread sequences, a group of Q pre-selected sequences can be used by the devices and this set is known by the receiver at the BS. As Q decreases, the probability of two different devices choosing the same spread sequence increases. This situation is called access collision and it can happen in a grant-free multiple access, such as in Carrier Sense Multiple Access (CSMA) employed in Wi-Fi networks [25].

However, while this collision in CSMA results in package drop, in MUSA it is not always fully destructive and data from two devices using the same spread sequence can still be successfully recovered. However, it is interesting to reduce this probability of collisions by increasing Q, at the cost of higher complexity on the BS receiver. Eq. (1) shows one MUSA codebook for $Q = 15$ and $L = 4$.

Since MUSA is a grant-free access scheme, a given device can transmit without previous scheduling from the BS. Each device randomly chooses a normalized code sequence $\mathbf{s}_k$, given by

$$\mathbf{s}_k = \frac{\mathbf{g}_q}{\sqrt{\mathbf{g}_q^{\mathrm{H}} \mathbf{g}_q}}, \tag{2}$$

to spread its data symbol before transmitting, where $\mathbf{g}_q$ is the qth column of $\mathbf{G}$. After the data symbol is spread, the number of resources necessary to transmit the information is equal the length of the code, L. Different modulation schemes can be used in this scenario. For instance, integration with Orthogonal Frequency Division Multiplexing (OFDM) is straightforward, since each time-frequency resource necessary to carry the L-long spread sequence can be seen as subcarrier of one specific OFDM block. It is worth mentioning that no pilot or training sequences are transmitted for channel estimation at the receiver side. Also, no information about the sequences chosen by the devices are transmitted. Fig. 2 shows the block diagram of the MUSA transmitter.

Initially, a block of U data bits $\mathbf{b}_k^{U \times 1}$ is generated by the kth device. This bit vector is encoded by a channel encoder with code rate $R = U/V$, resulting in the encoded block of V bits, $\mathbf{c}_k^{V \times 1}$. The encoded bits are mapped into in-pahse and quadrature symbols and, in this paper, Differential Binary Phase Shift Keying (DBPSK) has been chosen because it allows for phase tracking at the receiver side without previous knowledge of the CSI. Since mMTC devices transmit low amount of data per channel use, the low modulation order is adequate. The modulator introduces one reference symbol for phase tracking, resulting in a modulated block $\mathbf{d}_k^{L \times 1}$, where $L = V + 1$. The modulated block is spread by a tri-level complex sequence $\mathbf{s}_k^{J \times 1}$, arbitrary chosen from the codebook, resulting in the spread block $\mathbf{m}_k^{LJ \times 1}$, where J represents the length of the spread complex sequence. Next, the spreading block is allocated in a set of subcarriers from a OFDM symbol [18], which is generated by the Inverse Fast Fourier Transform (IFFT). It is assumed that each device receives $N < LJ$ subcarriers in each OFDM symbol. Therefore, LJ/N different OFDM symbols are needed to transmit the $\mathbf{m}_k$ block. The frame necessary to transmit the $\mathbf{m}_k$ block is represented by $\mathbf{X}_k^{N \times \frac{LJ}{N}}$.

III. Principles of the MUSA receiver

In the MUSA scheme, the devices share the OFDM time-frequency resources and the signal received in the nth subcarrier of the ith OFDM is given by

$$r^{n,i} = \sum_{k=1}^{K} h_k^{n,i} x_k^{n,i} + w^{n,i}, \tag{3}$$

where $n = 1, 2, \cdots, N$ and $i = 1, 2, \cdots, LJ/N$. The notation $x_k^{n,i}$ represents the value in the nth row and ith column of $\mathbf{X}_k$. The channel gain between the kth device and the BS at the nth subcarrier of the ith OFDM block is represented by $h_k^{n,i}$, while $w^{n,i}$ represents the Additive White Gaussian Noise (AWGN) noise samples with zero-mean and variance σ_w^2. The received samples can be organized in a matrix $\mathbf{R}^{N \times \frac{LJ}{N}}$ that contains the received MUSA frame.

Due to the low complexity required by mMTC devices, the close loop power control shall be avoided, since this feature increases the communication overhead and requires extra control processing at the IoT devices. Hence, the Signal-to-Interference-and-Noise Ratio (SINR) of the signal received from each device at the BS can vary significantly. This behavior is known as the near-far effect, which the closest device to the base station has, on average, the highest SINR, while the farthest one has the lowest SINR, on average, among the all other devices. This characteristic can be exploited by the receiver based on SIC. Furthermore, the difference among the SINRs from different devices can allow for successful data recovery when code collision from two different devices occurs.

The block diagram of a MUSA blind SIC-based MUD receiver is shown in Fig. 3. The SIC uses the recovered information received from a reliable link to remove the interference introduced in the information transmitted by other devices. This procedure is performed for all received data, starting with the signal with highest SINR to the signal with lowest SINR.

The matrix $\mathbf{R}^{N \times \frac{LJ}{N}}$ is defined after all LJ/N OFDM symbols are received. Assuming, without loss of generalization, that $x_1^{n,i}$ is the signal with highest SINR in the nth subcarrier of the ith OFDM symbol, then (3) can be rewritten as

$$r^{n,i} = \underbrace{h_1^{n,i} x_1^{n,i}}_{\text{desired signal}} + \underbrace{\sum_{k=2}^{K} h_k^{n,i} x_k^{n,i}}_{\text{interference}} + \underbrace{w^{n,i}}_{\text{noise}}. \tag{4}$$

It is expected that the high SINR of the first signal to be decoded and the powerful Forward Error Correction (FEC) guarantee that the information is received without error. In order to proceed with the SIC algorithm, the column-wise Fast Fourier Transform (FFT) is performed in $\mathbf{R}^{N \times \frac{LJ}{N}}$, decoupling the subcarriers. The resulting matrix is organized in a column vector $\hat{\mathbf{m}}_1^{LJ \times 1}$. The despread block employs cross-correlation between the received vector and the known spreading codes to identify the code sequence employed by the device with the highest SINR.

In [20], [21], [22], the authors omit the procedure to blindly estimate the complex spreading. In [22], the authors mention that there is a metric that can be used to estimate the spread sequence, however, no further details are presented. Since the spread sequence estimation is essential for recovering the data sent by the devices, this paper presents a detailed description on how these sequences can be estimated by the BS. In the procedure proposed in this paper, the BS's receiver uses all known spread sequences from $\mathbf{G}$ to recover the data from a given device. Each sequence will led to a version of de-spreaded signal. All resulting sequences are detected using the

$$\mathbf{G}=\begin{bmatrix} 1+j & -1-j & -1+j & -j & 1-j & 1 & -1-j & j & 1+j & -1+j & -1+j & -1 & 1+j & 1-j & 1-j \\ 1+j & j & 1+j & 1+j & -1-j & -1-j & j & 1-j & 1-j & 1+j & -j & 1 & -1+j & 1-j & -1+j \\ 1+j & j & -1+j & 1 & 1+j & -1-j & -1-j & 1-j & -1+j & 1-j & -1-j & 1-j & -1+j & -j & -1-j \\ 1 & -j & 1+j & -1+j & j & 1 & -1 & -j & -1 & -1-j & -1+j & 1+j & 1 & j & 1+j \end{bmatrix} \tag{1}$$

Fig. 2. Block diagram of MUSA transmitters.

Fig. 3. Block diagram of the blind detection SIC-based MUD receiver. The notation ∘ represents the Hadamard product.

DBPSK receiver, which does not require the CSI. After the DBPSK detection, the sequence is decoded by the FEC and the the Cyclic Redundancy Check (CRC) for each recovered sequence is verified. The spread sequence $\mathbf{s}_1$ that results in an error-free data is the sequence assumed to be employed by the device during transmission. Notice that this procedure requires that the received SINR must allow for the correct FEC decoding of the transmitted data. Hence, is it essential to select the signal from the device with highest SINR as the first one to be received. This approach has good performance when one device has a higher SINR than the others, as presented in [17]. If two or more devices have similar SINR, the system performance decreases significantly.

At the end of the described procedure, the information data block of the device with highest SINR can be canceled from the received signal $\mathbf{R}$ in order to proceed with the detection

of the information data block of the device with second highest SINR among all other remaining users. Therefore, the vector $\hat{\mathbf{b}}_1^{U\times 1}$, recovered by using $\mathbf{s}_1$, is encoded, modulated and spread. This signal is assumed to be equal to the one produced by the lth device and it must be subtracted from the received signal to cancel its interference on the other devices. The IFFT is applied to the feedback information block to generate the matrix $\hat{\mathbf{X}}_1^{N\times\frac{LJ}{N}}$. This signal must be weighted by the channel response in order to be subtracted from $\mathbf{R}^{N\times\frac{LJ}{N}}$. Although CSI knowledge is not necessary to demodulate the received sequence, this information is necessary for the SIC algorithm. $\hat{\mathbf{X}}_1^{N\times\frac{LJ}{N}}$ and $\mathbf{R}^{N\times\frac{LJ}{N}}$ are used to estimate the channel gain between the lth device and the BS.

As suggested in [21], the LS estimator is used to estimate the channel gain as expressed below:

$$\hat{\mathbf{h}}^i = \left(\mathbf{x}^{i\mathrm{H}}\mathbf{x}^i\right)^{-1}\mathbf{x}^{i\mathrm{H}}\mathbf{r}^i, \tag{5}$$

where $i = 1, 2, \cdots, LJ/N$, $\mathbf{x}^i$ and $\mathbf{r}^i$ denotes the vectors containing N points formed by the ith column of the $\hat{\mathbf{X}}_k$ and $\mathbf{R}$, respectively. LS estimation provides poor estimation of the channel gains since the presence of noise is ignored in the estimation process. Hence, LS estimation error propagates by the SIC receiver and compromises the performance of the MUD. According to [26], AI algorithms can be used to improve the overall channel estimation accuracy by reducing the LS estimation error and reducing the impact of noise that is dominant in low SNR region. Motivated by these advantages, an optimized DNN architecture is employed on top of the LS estimation, where the improved version of the LS channel estimation is used to remove the interference of the lth device from the received sequence. The details of the DNN used in to improve the LS estimation are presented next.

A. DNN-based channel estimator

Nowadays, DNN is used to mitigate or solve some problems in telecommunications area [27], [28], [29], in which it also includes the use of DNN applying in NOMA schemes [30], [31], [32], [33]. DNN is used in general to map input features to an output predicted values, through a set of mathematical operations performed by multiple connected layers, each containing multiple processing units called neurons. The DNN input-output mapping is accomplished by minimizing the DNN loss function that represents the input-output difference in the training phase, where the DNN is trained on known input-output data pairs. During the DNN training phase, the loss function minimization can be achieved through iteratively updating the DNN parameters denoted as weights to achieve the best possible performance. After that, the trained DNN performance is evaluated by its ability in predicting true outputs for new data inputs.

Let E be the number of DNN layers, with P neurons for each layer e, where $1 \leq e \leq E$ and $1 \leq p \leq P$. The neuron output $\tilde{y}_{e,p}$ can be expressed as follows:

$$\tilde{y}_{e,p} = f_{e,p}\left(\tilde{b}_{e,p} + \tilde{\omega}_{e,p}^{\mathrm{T}}\, \tilde{x}_e\right), \qquad (6)$$

where $\tilde{x}_e \in \mathbb{R}^{P_{e-1}\times 1}$, $\tilde{\omega}_{e,p} \in \mathbb{R}^{P_{e-1}\times 1}$, and $\tilde{b}_{e,p}$ represent the neuron input, weight vector, and bias respectively. Each neurons (e,p) performs a linear transformation represented by the activation function $f_{e,p}$ as shown in (6). Similarly, the DNN layer output can be expressed as follows

$$\tilde{y}_e = f_e\left(\tilde{b}_e + \tilde{W}_e^{\mathrm{T}}\, \tilde{x}_e\right), \; \tilde{x}_e = \tilde{y}_{e-1}. \qquad (7)$$

After defining the DNN architecture, DNN should be trained on known dataset in order to minimize the loss function that measures how far apart the predicted DNN outputs $\tilde{y}_E^P$ from the true outputs $\tilde{y}_E^{\mathrm{T}}$. The most common used loss function in regression problems is the mean squared error (MSE) function that can be defined as

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{train}}} \sum_{i=1}^{N_{\mathrm{train}}} \left\| \tilde{y}_E^P - \tilde{y}_E^{\mathrm{T}} \right\|^2, \qquad (8)$$

where N_{train} denotes the number of training samples. The main objective of the DNN is to update the total weights matrix $\tilde{W}$, where the MSE between the predicted and true DNN outputs is minimized. To do so, several DNN optimizers can be used, like the stochastic gradient descent, and adaptive moment estimation (ADAM) [34].

The proposed DNN-based channel estimation depends mainly on LS channel estimates $\hat{\mathbf{h}}^i$ obtained from (5), where DNN is employed as an additional non-linear processing unit on top of the LS estimation to correct the overall estimation error by learning the channel frequency domain characteristics, resulting in improved estimated channels.

The proposed DNN-based channel estimation proceeds in the following steps:

1) Initial LS estimation as shown in (5).
2) The obtained $\hat{\mathbf{h}}^i$ is transformed from complex-valued to real-valued domain by stacking the real and imaginary part vertically, such that $\hat{\mathbf{h}}_{\mathrm{R}}^i \in \mathbb{R}^{2N\times 1}$.
3) $\hat{\mathbf{h}}_{\mathrm{R}}^i$ is fed as an input to the proposed DNN.
4) The output of the DNN is transformed back to the complex-valued domain, such that $\hat{\mathbf{h}}_{\mathrm{DNN}}^i \in \mathbb{C}^{N\times 1}$.

The proposed DNN is trained using $N_{\mathrm{train}} = 8000$ training samples $(\hat{\mathbf{h}}^i, \mathbf{h}^i)$, where $\mathbf{h}^i$ denotes the perfect channel for the ith received symbol. The training dataset is generated using high SNR (40 dB), since the analysis provided in [26], show that training on high SNR values leads to a better DNN generalization functionality in lower SNRs due to the fact that in high SNR the impact of noise is low, thus the DNN is able to learn more the channel frequency domain characteristics. The mean squared error is chosen as a loss function that is optimized using ADAM optimizer. The testing datasets are generated using different SNRs = [4, 11.2, 18.4, 25.6, 32.8] dB, where the trained DNN is employed to evaluate the proposed DNN-based estimator. Moreover, intensive simulations were

TABLE I
PROPOSED DNN PARAMETERS.

(Hidden layers; Neurons per layer)	(1;24)
Activation function	ReLU
Number of epochs	500
Training samples	8000
Testing samples	2000
Batch size	128
Optimizer	ADAM
Loss function	MSE
Learning rate	0.001
Training SNR	40 dB
Testing SNRs	[4, 11.2, 18.4, 25.6, 32.8] dB

performed on several DNN architectures in order to select the more efficient one in terms of performance and complexity. The simulations show that employing one hidden layer DNN with N neurons is sufficient. Table I provides the parameters of the proposed DNN-based estimator.

IV. NUMERICAL RESULTS

The BLER is the main key performance indicator (KPI) for MUSA performance evaluation. In this paper, each device encodes U data bits using a Polar Code, creating a codeword with length V. A reference bit for the DBPSK modulation is added to the codewords, resulting in a block of $L = V + 1$ bits, which is transmitted by the devices using OFDM symbols with N subcarriers, as described in Section II. Each device can spread their data using one of the Q complex spread sequences, each one with length J. The sequences employed in this paper have been chosen to provide acceptable auto-correlation properties, allowing the BS to blindly estimate the sequences employed by the devices. On the receiver side, the blocks are recovered by the BS as described in Section III. After MUSA MUD, the codewords are processed by the Polar decoder. The codewords that cannot be successfully corrected are accounted for the BLER.

Table II brings the parameters used for the simulations presented in this section. Since the length of the spreading sequence is $J = 4$, the spreading block has $LJ = 192 \cdot 4 = 768$ elements. These elements are mapped into $LJ/N = 64$ OFDM symbols with $N = 12$ subcarriers in each one. Therefore, one MUSA frame occupies the same Physical Resource Block (PRB) during 9 time-slots in a time-frequency grid, following the LTE structure.

TABLE II
SIMULATION PARAMETERS

Parameter	Value	Parameter	Value
U	64	J	4
V	191	N	12
L	192	Q	64

A time-varying frequency-selective channel based on the Tapped Delay Line (TDL)-D [35] was used in the simulations, assuming independent block fading for each PRB. Finally, it

is assumed that all users transmit a data block in all MUSA frame. The device with the highest SNR achieves 40 dB, while the lowest SNR is 4 dB. The SNR for the devices is uniformly distributed between to these upper and lower limits.

Fig. 4 shows the BLER performance for the proposed MUSA framework under different overload values, defined as $\lambda = K/J \cdot 100\%$. Also, both AWGN and TDL-D channels have been considered. The perfect channel estimation is denoted MUSA-ideal estimation, while MUSA-LS and MUSA-DNN are used to denote when LS and DNN-based estimation were employed. The acceptable BLER target is 10%.
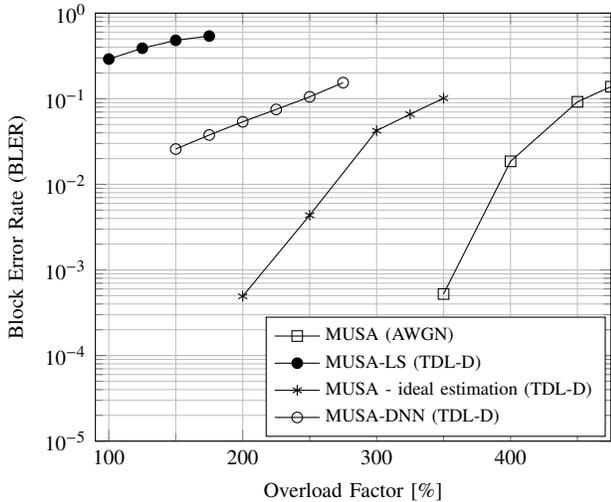
Fig. 4. BLER performance of MUSA scheme under different overload values.

It is possible to observe in Fig. 4 that MUSA has good BLER performance over AWGN channel, since the channel response is flat and no frequency-response estimation is required. Hence, SIC algorithms does not suffer with error propagation and high overloading factors are possible for the target BLER, allowing λ up to 450%.

When the TDL-D channel is considered, Fig. 4 shows that the MUSA BLER performance decreases, even when CSI is assumed to be available at the BS. In this case, the maximum overloading factor that can be achieved at the target BLER is 350%. If the LS estimator is employed, the MUSA BLER performance becomes unacceptable and the system is considered to be inoperative even for overloading factors below 100%. The DNN-based estimation significantly improves the performance of the MUSA scheme over TDL-D channels. The improved channel estimation provided by the DNN reduced the error propagation and allows for the SIC algorithm to properly cancel the interference among the devices. Applying DNN as a post processing unit on top of the LS estimation leads to a significant normalized mean squared error (NMSE) performance improvement, as shown in Fig. 5 thus, allowing the MUSA to achieve an overload factor of 250% over TDL-D channel, which means that up to 10 users can share the 4 PRB of the MUSA frame at the target BLER. This means that the DNN-based estimation brought the MUSA system from the inoperative state to a situation where an considerable spectrum efficiency gain can be achieved. The use of DNN
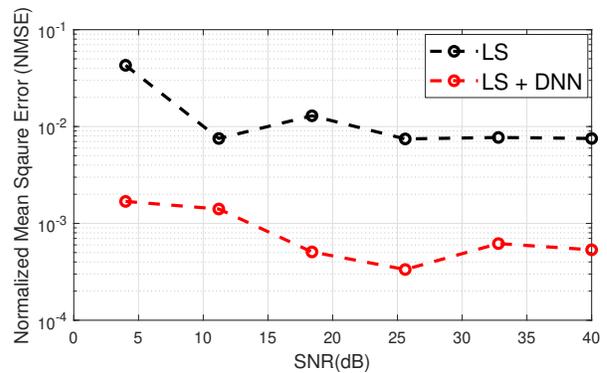
Fig. 5. NMSE performance of the proposed DNN-based estimator. This figure is better seen in colored mode.

in other processes in the receiver chain can further improve the overall MUSA system performance, bringing it closer to the performance achieved with perfect channel estimation, and this will be the topic of future research efforts.

Fig. 6 shows the performance of MUSA-DNN for different SNR conditions. In this case, $\text{SNR}_a = a\frac{E_b}{N_0}$ is used in the simulations.
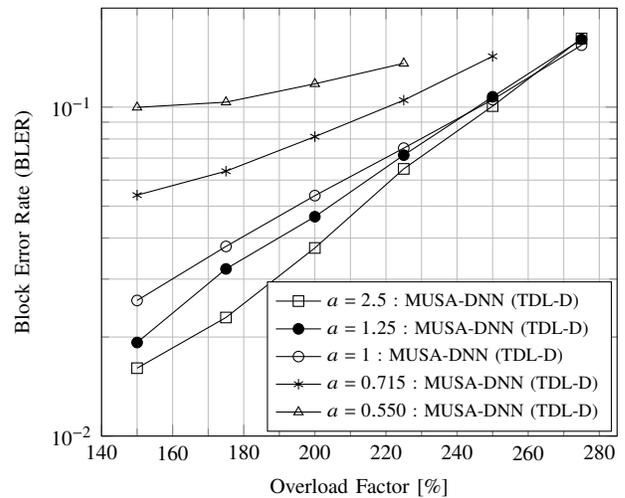
Fig. 6. BLER performance of MUSA scheme under different values of overload and SNR.

Notice that, as expected, by decreasing the SNR, the MUSA-DNN performance also decreases, resulting in smaller overladiing factor for the target BLER. It is possible to notice that, for $a = 0.715$ and $a = 0.550$, the achieved overload factors are 200% and 150%, respectively. In other words, when the SNR decreases, the system needs to allocate fewer users in the same transmissions resource in order to achieve the BLER target, decreasing the throughput in the system. On the other hand, increasing the SNR increases the BLER performance curves, but it does not improve the overload factor at the target BLER. For $a > 1$, the maximum overload factor is 250% for a BLER of 10^{-1}.

The impact of the code rate R on the MUSA performance was also evaluated. In this scenario $V = 191$ and different values of U are used, changing the the code rate. Fig. 7 depicts the

MUSA performance in terms of BLER for $U \in [64, 32, 16]$. As expected, when the code rate is decreased, better BLER performance is achieved because of the higher error correction capabilities of the Polar Code, at the cost of lower throughput per device. With a $R = 16/191 = 0.08377$, it is possible to allocate 12 users transmitting their data in the same elements and, even so, the system achieves a BLER less than 10%, but a data rate reduction of 4 times when compared with the scenario presented in Fig. 4.
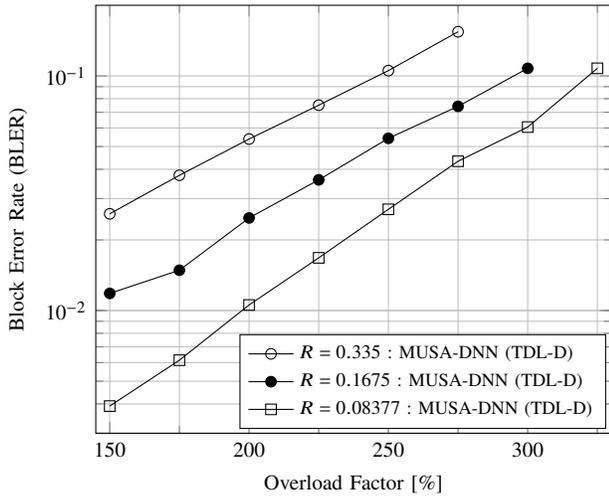
Fig. 7. BLER performance of MUSA scheme under different values of overload and code rate.

Finally, a comparison between the performances of the MUSA and Orthogonal Frequency Division Multiple Access (OFDMA) is analyzed below. It is worth to mention that OFDMA is a orthogonal multiple access scheme. Fig. 8 illustrates the main difference between the time-frequency resource allocation for MUSA and OFDMA. While MUSA uses the overload factor and allows for all K users to transmit their data in all time-frequency resources, the OFDMA divides the time-frequency resources among the K users in an orthogonal manner. In this paper, a block containing 64 OFDM symbols, each one with $N = 12$ subcarriers was defined to compare the performance of the two schemes.

In MUSA, the K devices spread their data to the entire block. In the OFDMA, an device must transmit the data on specific positions of the time-frequency grid. Furthermore, the devices in the OFDM scheme must employ pilot subcarriers to estimate the channel response and must follow the synchorinization procedure and reource allocation coordinated by the BS. These processes requires high processing capabilities and also demands power allocation for synchronization and channel estimation signalling, reducing the power and spectrum efficiencies of the system.

Fig. 9 shows the BLER for MUSA-DNN and OFDM with 6 and 8 users for different values of SNR, obtained by varying a. Because OFDMA does not suffer from multi-user interference, its performance improves as the SNR increases. The MUSA, on the other hand, suffers a error floor due to the multi-user interference. However, it is interesting to observe that for the
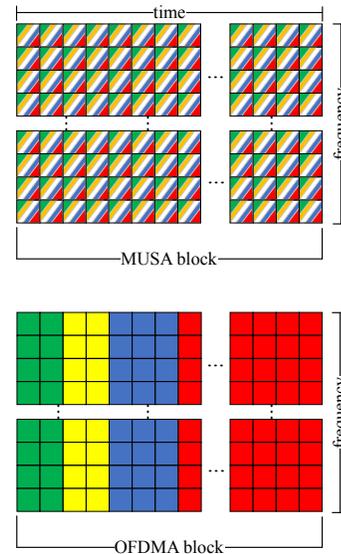
Fig. 8. Resource allocation for MUSA and OFDMA. Colors represent the information of different devices.

target BLER of 10^{-1}, both schmes have similar performance, requiring $a = 0.55$ to achieve this KPI.
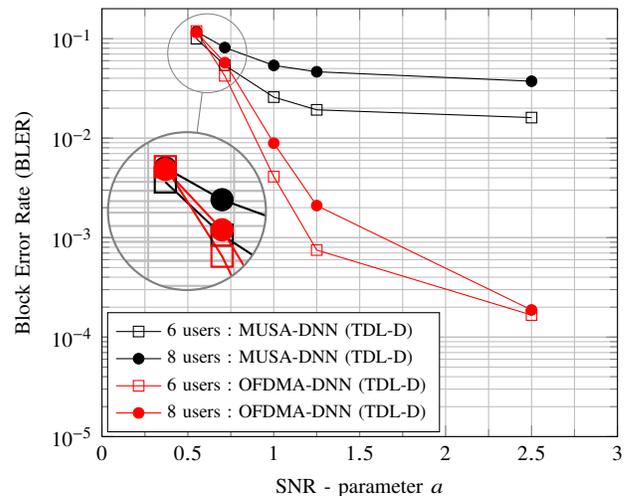
Fig. 9. BLER performance of MUSA and OFDMA schemes under different values of overload and SNR. This figure is better seen in colored mode.

V. CONCLUSION

The advent of IoT applications will require a massive number of power-limited devices to be connected in the mobile network. NOMA techniques can provide the flexibility to increase the spectrum efficiency without increasing the complexity of these devices and MUSA is an interesting candidate. Previous papers has shown that MUSA can achieve very high overloading factor without including any information in the transmitted data that aids the channel estimation and the estimation of the spread sequences employed by the device. However, the procedures performed by the receiver in the BS were not detailed. In this paper, a complete framework for the MUSA is described. It has been shown that the MUSA

JOURNAL OF COMMUNICATION AND INFORMATION SYSTEMS, VOL. 36, NO.1, 2021.

126

performance can be severely penalized by doubly dispersive channels. This paper has also shown that the use of DNN benefits the quality of the the channel estimation, improving the overall system performance and allowing MUSA to achieve reasonable an overloading factor. AI algorithms can be used in other processes in the receiving chain, which might improve even further the BLER performance or reduce the system complexity on the BS side, making the MUSA an interesting candidate for future IoT application over mobile networks.

REFERENCES

[1] H. Yin, L. Zhang, and S. Roy, "Multiplexing URLLC traffic within eMBB services in 5G NR: Fair scheduling," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 1080–1093, 2021, doi: 10.1109/TCOMM.2020.3035582.

[2] A. Anand, G. de Veciana, and S. Shakkottai, "Joint scheduling of URLLC and eMBB traffic in 5G wireless networks," *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 477–490, 2020, doi:10.1109/TNET.2020.2968373.

[3] L. L. Mendes and et al, "Enhanced Remote Areas Communications: The missing scenario for 5G and beyond 5G networks," *IEEE Access*, vol. 8, pp. 219 859–219 880, 2020, doi: 10.1109/ACCESS.2020.3042437.

[4] S. R. Pokhrel, J. Ding, J. Park, O.-S. Park, and J. Choi, "Towards enabling critical mMTC: A review of URLLC within mMTC," *IEEE Access*, vol. 8, pp. 131 796–131 813, 2020, doi: 10.1109/ACCESS.2020.3010271.

[5] V. Gupta, S. K. Devar, N. H. Kumar, and K. P. Bagadi, "Modelling of IoT traffic and its impact on LoRaWAN," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6, doi:10.1109/GLOCOM.2017.8254512.

[6] M. Gharbieh, H. ElSawy, H. Yang, A. Bader, and M. Alouini, "Spatiotemporal Model for Uplink IoT traffic: Scheduling and random access paradox," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8357–8372, 2018, doi:10.1109/TWC.2018.2876522.

[7] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014, doi:10.1109/MCOM.2014.6815890.

[8] O. O. Oyerinde, "Comparative study of overloaded and underloaded NOMA schemes with two multiuser detectors," in *2019 IEEE 2nd Wireless Africa Conference (WAC)*, 2019, pp. 1–5, doi: 10.1109/AFRICA.2019.8843410.

[9] G. P. Aquino and L. L. Mendes, "Sparse code multiple access on the generalized frequency division multiplexing," *EURASIP Journal on Wireless Communications and Networking*, vol. 20, 2020, doi: 10.1186/s13638-020-01832-z. [Online]. Available: https://doi.org/10.1186/s13638-020-01832-z

[10] B. Wang, K. Wang, Z. Lu, T. Xie, and J. Quan, "Comparison study of non-orthogonal multiple access schemes for 5G," in *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, June 2015, pp. 1–5, doi:10.1109/BMSB.2015.7177186.

[11] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2017, doi:10.1109/COMST.2016.2621116.

[12] L. Dai, B. Wang, Y. Yuan, S. Han, C. l. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, September 2015, doi:10.1109/MCOM.2015.7263349.

[13] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, September 2013, pp. 332–336, doi:10.1109/PIMRC.2013.6666156.

[14] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access (PDMA) - a novel non-orthogonal multiple access for 5G radio networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, p. 3185–3196, 2017, doi:10.1109/TVT.2016.2596438.

[15] Y. Chen, F. Schaich, and T. Wild, "Multiple access and waveforms for 5G: IDMA and universal filtered multi-carrier," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–5, doi:10.1109/VTCSpring.2014.7022995.

[16] D. Fang, Y. Huang, Z. Ding, G. Geraci, S. Shieh, and H. Claussen, "Lattice partition multiple access: A new method of downlink non-orthogonal multiuser transmissions," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6, doi:10.1109/GLOCOM.2016.7841947.

[17] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-user shared access for Internet of Things," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5, doi: 10.1109/VTCSpring.2016.7504361.

[18] ZTE, "3GPP R1-162226: Discussion on multiple access for new radio interface," 3GPP, Tech. Rep., 2016.

[19] Y. Chunlin, Y. Zhifeng, L. Weimin, and Y. Yifei, "Non-orthogonal multiple access schemes for 5G," *ZTE Communications*, vol. 14, no. 4, pp. 11–16, October 2016, doi:10.3969/j.

[20] ZTE, "3GPP R1-164270: Receiver implementation for MUSA," 3GPP, Tech. Rep., 2016.

[21] ——, "3GPP R1-166404: Receiver details and link performance for MUSA," 3GPP, Tech. Rep., 2016.

[22] E. M. Eid, M. M. Fouda, A. S. T. Eldien, and M. M. Tantawy, "Performance analysis of MUSA with different spreading codes using ordered SIC methods," in *2017 12th International Conference on Computer Engineering and Systems (ICCES)*, Cairo, Egypt, December 2017, pp. 101–106, doi:10.1109/ICCES.2017.8275286.

[23] J. Qi, J. Du, S. M. Siniscalchi, X. Ma, and C.-H. Lee, "Analyzing upper bounds on mean absolute errors for deep neural network-based vector-to-vector regression," *IEEE Trans. Signal Process.*, vol. 68, pp. 3411–3422, 2020, doi: 10.1109/TSP.2020.2993164.

[24] M. Majmundar, N. Sandhu, and J. Reed, "Adaptive single-user receivers for direct-sequence spread-spectrum CDMA systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 379–389, 2000, doi: 10.1109/25.832968.

[25] Y. Imagaki, K. Kashiki, K. Yamazaki, and A. Yamaguchi, "Novel Wi-Fi Throughput Estimation Method Considering CSMA/CA Behavior," in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, 2012, pp. 1–5, doi: 10.1109/VETECS.2012.6240239.

[26] A. K. Gizzini, M. Chafii, A. Nimr, and G. Fettweis, "Enhancing least square channel estimation using deep learning," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020, pp. 1–5, doi:10.1109/VTC2020-Spring48590.2020.9128890.

[27] C.-C. Kao, C.-W. Chang, C.-P. Cho, and J.-Y. Shun, "Deep learning and ensemble learning for traffic load prediction in real network," in *2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE)*, 2020, pp. 36–39, doi: 10.1109/ECICE50847.2020.9302005.

[28] S. Sudhakaran, A. Venkatagiri, P. A. Taukari, A. Jeganathan, and P. Muthuchidambaranathan, "Metropolitan cellular traffic prediction using deep learning techniques," in *2020 IEEE International Conference on Communication, Networks and Satellite (Comnetsat)*, 2020, pp. 6–11, doi: 10.1109/Comnetsat50391.2020.9328937.

[29] A. Al-Baidhani and H. H. Fan, "Learning for detection: A deep learning wireless communication receiver over Rayleigh fading channels," in *2019 International Conference on Computing, Networking and Communications (ICNC)*, 2019, pp. 6–10, doi: 10.1109/ICCNC.2019.8685517.

[30] N. Ye, X. Li, H. Yu, A. Wang, W. Liu, and X. Hou, "Deep learning aided grant-free NOMA toward reliable low-latency access in Tactile Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2995–3005, 2019, doi: 10.1109/TII.2019.2895086.

[31] A. Emır, F. Kara, and H. Kaya, "Deep learning-based joint symbol detection for NOMA," in *2019 27th Signal Processing and Communications Applications Conference (SIU)*, 2019, pp. 1–4, doi: 10.1109/SIU.2019.8806600.

[32] M. K. Hasan, M. Shahjalal, M. M. Islam, M. M. Alam, M. F. Ahmed, and Y. M. Jang, "The role of deep learning in NOMA for 5G and beyond communications," in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2020, pp. 303–307, doi: 10.1109/ICAIIC48513.2020.9065219.

[33] K. Fathimath Shamna, C. Ismayil Siyad, S. Tamilselven, and M. K. Manoj, "Deep learning aided NOMA for user fairness in 5G," in *2020 7th International Conference on Smart Structures and Systems (ICSSS)*, 2020, pp. 1–6, doi: 10.1109/ICSSS49621.2020.9202308.

[34] S. De, A. Mukherjee, and E. Ullah, "Convergence guarantees for RMSProp and ADAM in non-convex optimization and an empirical comparison to Nesterov acceleration," *arXiv: Learning*, 2018, available in: https://arxiv.org/abs/1807.06766v3.

[35] ETSI, "3GPP TR38.900: Study on channel model for frequency spectrum above 6 GHz. version 14.2.0 release 14." ETSI, Tech. Rep., 2017.

Guilherme Pedro Aquino has a Ph.D. in Electrical Engineering from the Federal University of Itajubá, UNIFEI. Currently, he is with the National Institute of Telecommunications, INATEL, as a professor and researcher at the Reference Center for Radio-communications, conducting researches in non-orthogonal multiple access and efficient cooperative spectrum sensing strategies.

Luciano Leonel Mendes received the B.Sc. and M.Sc. in Electrical Engineer from Inatel, Brazil in 2001 and 2003, respectively. In 2007 he received the Doctor in Electrical Engineer degree from Unicamp, Brazil. Since 2001 he is a professor at Inatel, Brazil. He has coordinated the Master Program at Inatel and several research projects funded by FAPEMIG, FINEP and BNDES. He has been a visiting researcher at Vodafone Chair Mobile Communications Systems at the TU Dresden between 2013 and 2015. His main area of research is waveforms for 5G networks and future mobile communication systems.

Tiago Cardoso has MSc. in Microelectronics Engineering from Federal University of Itajuba, UNIFEI, 2010. Since 2009 he is a researcher and designer at National Institute of Telecommunications, INATEL, where develop algorithms in VHDL, Verilog, Python, and C++ for wireless communications. These algorithms are implemented in Softwares Defined Radios that includes 5G waveforms like FOFDM/GFDM, error-correcting codes, like Polar Codes, and digital power amplifiers linearization for LTE.

Marwa Chafii received her Ph.D. degree in Telecommunications in 2016, and her Master's degree in the field of advanced wireless communication systems (SAR) in 2013, both from CentraleSupélec, France. Between 2014 and 2016, she has been a visiting researcher at Poznan University of Technology (Poland), University of York (UK), Yokohama National University (Japan), and University of Oxford (UK). She joined the Vodafone Chair Mobile Communication Systems at the Technical University of Dresden, Germany, in February 2018 as a research group leader. Since September 2018, she is research projects lead at Women in AI and an associate professor at ENSEA, France, where she holds a Chair of Excellence on Artificial Intelligence from CY Initiative. She received the prize of the best Ph.D. in France in the field of Signal, Image & Vision, and she has been nominated in the top 10 Rising Stars in Computer Networking and Communications by N2Women in 2020. Her research interests include machine learning for wireless communications, advanced waveform design, indoor localization and channel estimation. Since 2019, she serves as Associate Editor at IEEE Communications Letters where she received the Best Editor Award in 2020. She is currently vice-chair of the IEEE ComSoc Emerging Technology Initiative on Machine Learning for Communications and managing the Gender Committee of the AI4EU community.

Abdul Karim Gizzini received the bachelor degree in computer and communication engineering (B.E) from the IUL university of Lebanon, in 2015 followed by the M.E degree in 2017. His master thesis was hosted by the Lebanese national council for scientific research (CNRS). Since 2019 he is pursuing the Ph.D. degree in wireless communication engineering in ETIS laboratory which is a joint research unit at CNRS (UMR 8051), ENSEA, and CY Cergy Paris University in France. His Ph.D. thesis focuses mainly on deep learning based channel estimation in high mobility vehicular scenarios.