State University of Londrina
Center of Technology and Urbanism
Department of Electrical Engineering
Master's Program in Electrical Engineering

João Henrique Inacio de Souza

Artificial Intelligence Techniques to Enable Massive Connectivity and XL-MIMO in the Future Wireless Communication Networks

Londrina
2021

João Henrique Inacio de Souza

Artificial Intelligence Techniques to Enable Massive Connectivity and XL-MIMO in the Future Wireless Communication Networks

Dissertation presented to the Graduate Program in Electrical Engineering at the State University of Londrina for obtaining the degree of Master in Electrical Engineering.

Concentration area:
Telecommunications Systems

Supervisor:
Prof. Dr. Taufik Abrão

Londrina
2021

Souza, João Henrique Inacio de

Artificial Intelligence Techniques to Enable Massive Connectivity and XL-MIMO in the Future Wireless Communication Networks. Londrina, 2021. 83 p.

Supervisor: Prof. Dr. Taufik Abrão

Dissertation (Master) – Department of Electrical Engineering – State University of Londrina

1. Artificial intelligence, 2. Massive multiple-input multiple-output (MIMO), 3. Machine-type communications, 4. Deep learning.

João Henrique Inacio de Souza

Artificial Intelligence Techniques to Enable Massive Connectivity and XL-MIMO in the Future Wireless Communication Networks

Dissertation presented to the Graduate Program in Electrical Engineering at the State University of Londrina for obtaining the degree of Master in Electrical Engineering.

Approved on June 25, 2021.

Examining Board

Prof. Dr. Taufik Abrão
Department of Electrical Engineering
State University of Londrina
Supervisor

Prof. Dr. José Carlos Marinello Filho
Department of Electrical Engineering
Federal University of Technology – Paraná

Prof. Dr. Sylvio Barbon Junior
Department of Computer Science
State University of Londrina

Prof. Dr. Richard Demo Souza
Department of Electrical and Electronics Engineering
Federal University of Santa Catarina

Londrina
2021

Acknowledgements

First of all, I thank my parents Vera and José for encouraging me in my choices. Also, I wish to thank my supervisor Prof. Taufik Abrão for giving me the opportunity to conduct the Master's studies. Finally, I thank my family and friends for their kindness, especially in this difficult time that we have been living since the beginning of 2020.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 141445/2020-3.

Resumo

As redes de comunicação sem-fio enfrentam um aumento exponencial da demanda dos usuários, além de um rápido crescimento no número de dispositivos conectados, impulsionado principalmente pela revolução da *Internet* das coisas. Visando atender à segunda necessidade, as redes sem-fio da quinta geração (5G) implementarão o serviço *masssive machine-type communications* (mMTC), projetado para fornecer conectividade massiva para dispositivos que possuem atividade esporádica, são equipados com *hardware* simples, e são alimentados por baterias de baixa capacidade. Por outro lado, as crescentes demandas por maior capacidade e taxa de dados serão assistidas por estações rádio-base equipadas com transceptores de múltiplas antenas. Atender a estes cenários é uma tarefa desafiadora devido à escassez de recursos de rádio combinada com as dificuldades técnicas enumeradas a seguir. Os esquemas de acesso à rede convencionais são extremamente ineficientes ao lidar com grandes números de tentativas de conexão. Ao mesmo tempo, estações rádio-base equipadas com grandes números de antenas apresentam gargalos de implementação decorrentes da alta largura de banda de interconexão e da complexidade computacional. Considerando os desafios mencionados, nesta Dissertação de Mestrado nós investigamos protocolos de acesso aleatório e o transceptor *extra-large scale massive* MIMO (XL-MIMO) para viabilizar, respectivamente, a conectividade massiva no serviço mMTC e a implementação de transceptores de múltiplas antenas. Adicionalmente, examinamos técnicas de inteligência artificial para projetar protocolos de acesso aleatório livres de concessão e algoritmos de alocação de recursos. Nossas contribuições são enumeradas a seguir. Propomos protocolos de acesso livres de concessão baseados nas redes neurais convolucionais que atingem alta performance com baixa complexidade computational na detecção de dispositivos ativos. Ao mesmo tempo, elaboramos uma análise compreensiva dos protocolos de acesso livres de concessão disponíveis na literatura. Em se tratando do transceptor XL-MIMO, propomos procedimentos distribuídos e centralizados baseando-se nos algoritmos genéticos para, conjuntamente, selecionar as antenas ativas e alocar potência visando maximizar a eficiência espectral do sistema. Os procedimentos propostos são adequados para implementações com número limitado de transceptores de rádio-frequência, atingindo largura de banda de interconexão e complexidade computacional reduzidas, quando comparados com os métodos de referência.

Palavras-chave: Inteligência artificial, *massive multiple-input multiple-output* (MIMO), *machine-type communications*, *deep learning*.

Abstract

The wireless communication networks face an exponential increase in users' demand, as well as a rapid growth in the number of connected devices, driven principally by the revolution of the Internet of things (IoT). Motivated by the second need, the wireless networks of the fifth generation (5G) will implement the massive machine-type communications (mMTC) service, designed to provide massive connectivity for devices with sporadic activity, with simple hardware, and powered by low-capacity batteries. On the other hand, the increasing demands for higher capacity and data rates will be fulfilled by multi-antenna base-station (BS) transceivers. Serving these scenarios is challenging due to the scarcity of radio resources along with the technical difficulties in the following. The conventional network access schemes are extremely inefficient for handling massive numbers of connection attempts. In addition, BSs equipped with a massive number of antennas present the bottlenecks of high interconnection bandwidth and computational complexity. Considering the mentioned challenges, in this Master's Dissertation we investigate random access protocols and the extra-large scale massive MIMO (XL-MIMO) transceiver to enable, respectively, massive connectivity on the mMTC service and the deployment of massive multi-antenna transceivers. We examine artificial intelligence techniques to design grant-free random access protocols and resource allocation algorithms. Our contributions are as follows. We propose grant-free access protocols based on the convolutional neural network with high performance and low complexity in detecting the set of active devices. At the same time, we conceive a comprehensive analysis of the grant-free access protocols in the literature. On the matter of the XL-MIMO transceiver, we propose distributed and centralized procedures based on the genetic algorithms for joint antenna selection and power allocation to maximize the spectral efficiency of the system. The proposed procedures are suitable for array implementations with limited number of radio-frequency transceivers, while attaining low interconnection bandwidth and complexity, when compared with the reference methods.

Keywords: Artificial intelligence, massive multiple-input multiple-output (MIMO), machine-type communications, deep learning.

List of Acronyms

5G	Wireless networks of the fifth generation
ADAM	Adaptive moment
AI	Artificial intelligence
AMP	Approximate message passing
AS	Antenna selection
BS	Base-station
CD	Coordinate descent
CNN	Convolutional neural network
COVID-19	Coronavirus disease 2019
CPU	Central processing unit
CSI	Channel state information
DET	Detection error trade-off
DFN	Deep feedforward neural network
DGA-RA	Quasi-distributed genetic algorithm for resource allocation
DL	Deep learning
E2E	End-to-end
eMBB	Enhanced mobile broadband
flops	Floating point operations per second
GA	Genetic algorithm
GA-RA	Genetic algorithm for resource allocation
IoT	Internet of things
LASSO	Least absolute shrinkage and selection operator
M2M	Machine-to-machine
MAC	Medium access control

MIMO	Multiple-input multiple-output
MMSE	Minimum mean-squared error
mMTC	Massive machine-type communications
MMV	Multiple-measurement vector
MSE	Mean-squared error
N-AS	Norm antenna selection
NMSE	Normalized mean-squared error
NN	Neural network
NNLS	Non-negative least-squares
PA	Power allocation
QoS	Quality-of-service
RA	Resource allocation
ReLU	Rectified linear unit
ROC	Receiver operating characteristic
RPU	Remote processing unit
SCMAX-AS	Sum-capacity maximization antenna selection
SE	Spectral efficiency
SGD	Stochastic gradient descent
URLLC	Ultra-reliable low-latency communications
VAMP	Vector approximate message passing
XL-MIMO	Extra-large scale massive multiple-input multiple-output
ZF	Zero-forcing

List of Notations

$\mathbf{a}$	lowercase bold letters represent column vectors;
$\mathbf{A}$	bold capital letters represent matrices;
$\mathcal{A}$	calligraphic capital letters represent finite sets;
$\mathbf{I}_n$	identity matrix of size n;
$\{\cdot\}^T$	transpose operator;
$\{\cdot\}^H$	conjugate transpose operator;
$\{\cdot\}^{-1}$	inverse matrix operator;
$\mathbb{R}$	set of the real numbers;
$\mathbb{Z}_+$	set of the non-negative integer numbers;
$\mathbb{E}[\cdot]$	expectation operator;
$\mathcal{CN}(\mu, \sigma^2)$	circularly symmetric complex Gaussian distribution with mean μ and variance σ^2;
Bernoulli(p)	Bernoulli distribution with mean p;
$f_i \circ f_j$	composition of the functions f_i and f_j;
$\|\cdot\|_p$	ℓ_p norm operator;
$\tanh(\cdot)$	hyperbolic tangent function.

List of Symbols

$\mathbf{r}_0$	input of a neural network model;
$\mathbf{r}_\ell$	output of the ℓ-th layer of a neural network model;
$f_\ell(\cdot,\cdot)$	function which represents the transformation of the ℓ-th layer of a neural network model;
θ_ℓ	set of parameters of the ℓ-th layer of a neural network model;
L'	depth of a neural network model;
$\boldsymbol{\theta}$	set with all the parameters of a neural network model;
$\sigma(\cdot)$	activation function;
$\mathbf{W}_\ell$	weights of the ℓ-th densely connected layer of a neural network model;
$\mathbf{b}_\ell$	bias vector of the ℓ-th densely connected layer of a neural network model;
B	number of subarrays at the base-station;
N_b	number of radio frequency transceivers at the b-th subarray;
M_b	number of antennas at the b-th subarray;
p_k	allocated power for the k-th user;
D_m	state indicator of the m-th antenna;
$P_{\max}$	maximum power constraint;
σ_z^2	thermal noise power;
$\mathbf{P}$	matrix with the allocated powers;
$\mathbf{D}$	matrix with the antenna state indicators;
$\mathbf{H}$	channel matrix;
$\mathcal{M}_b$	set of antennas at the b-th subarray;
M	number of antennas at the base-station;
K	number of users/devices in the cellular system;
L	length of the preambles/pilots in symbols;
N_n	number of neurons in the hidden layers of the DFN;
N_f	number of feature maps in the hidden layers of the CNN;
N_w	length of the filters in the layers of the CNN;
τ	threshold of the hard decision module;
Δ	undersampling ratio.

List of Figures

Contents

1 Introduction

The wireless networks face an exponential increase in users' demand, as well as the emergency of applications with high potential of social impact in healthcare, education, security, transportation, and industry. The **Internet of things** (IoT) promise to transform our everyday life with more efficiency, comfort, and safety connecting processes, tools, wearable devices, and sensors to the Internet [1]. Following this trend, the **machine-to-machine** (M2M) connections will be one of the primary contributors for the mobile traffic growth. In 2018, the M2M connections represented 13% of the 5.1 billion global mobile connections. Such number is predicted to grow rapidly, representing 34% of the 5.7 billions mobile connections by 2023 [2]. The challenges carried by these demands are supporting thousand of devices in a single cell with scalable and flexible network architectures, achieving high energy efficiency to save battery lifetime, providing joint low-latency and extreme-high reliability, and providing high data rates with high availability. The wireless networks of the next generation must implement technologies, network designs, and algorithms in order to meet the requirements of the emerging applications, as well as the increased users' demand, ensuring affordable financial costs and low environmental impacts.

The recent episode of the **coronavirus disease 2019 (COVID-19) pandemic** have been exposed the paramount need of robust communication networks, as it have resulted in a drastic increase on the Internet traffic. The strict social distancing measures imposed by many countries have produced growths on the remote working, education, e-commerce, and entertainment, such as online gaming and video streaming. After the COVID-19 pandemic started, a global provider of web and Internet security services has registered, just in one month, an unprecedented growth of about 30% in global Internet traffic [3]. As a response to the increasing demands, mobile-access regulators across the world have granted temporary spectrum licenses, released temporary bands, and implemented more flexible directives to use the licensed spectrum [4]. At the same time, major streaming providers have adjusted the quality levels of the distributed media aiming to optimize the bandwidth consumption. The COVID-19 pandemic has been resulted in possibly **permanent changes** in the Internet usage pattern, and also has been forced or accelerated the digital transformation of many enterprises. The wireless network infrastructure must follow this critical changes by delivering scalable, flexible, and resilient architectures.

Considering the rapid growth of the wireless mobile networks usage, it is necessary to examine the emerging application cases, and identify their requirements and bottlenecks in order to propose suitable techniques to fulfill the technical limitations. In this Master's Dissertation, we investigate the typical use case of the **massive machine-type com-**

munications (mMTC) service, analyzing deeply different types of **grant-free random access protocols** to handle large groups of machine-type devices. On this matter, we propose grant-free random access protocols based on machine learning. Also, we investigate the **extra-large scale massive MIMO** (XL-MIMO) transceiver, which is a strong candidate to handle crowded communication scenarios. We study the resource allocation (RA) problem of joint antenna selection (AS) and power allocation (PA) in XL-MIMO arrays, proposing distributed and centralized solutions based on the genetic algorithms (GAs). One common point of these studies is the presence of **artificial intelligence** (AI) techniques, such as machine learning and GAs. These techniques have been considered heavily in works on wireless communications in order to solve intricate problems in an efficient way. Especially, machine learning techniques are known to address scenarios with cumbersome mathematical models [5]. Our contributions are presented by means of three scientific publications (listed in Section 1.6) addressing the listed issues. In the following, we provide an overview of the fundamentals applied during the development of these works.

1.1 5G Wireless Communications

The fifth generation (5G) wireless networks are the evolution of the broadband cellular networks designed as a response to the emerging applications and the increased users' demand. The evolution to 5G is based on providing massive connectivity, improving energy efficiency, capacity and data rate, as well as reducing the cost and end-to-end (E2E) latency. The main technical objectives pursued in designing these networks are [6]:

(a) Increase data volume in 1000 times;

(b) Increase user data rate in 10 to 100 times;

(c) Increase the number of connected devices in 10 to 100 times;

(d) Extend the battery lifetime in 10 times;

(e) Reduce the E2E latency in 5 times.

The 5G wireless networks are the necessary evolution to fulfill the increasing demands of a connected world. However, these disruptive demands result in fundamental technical difficulties which need intensive effort to be met.

The 5G networks will assist heterogeneous services assigned to different applications with specific requirements. The main three operation modes of 5G are the *enhanced mobile broadband* (eMBB), *ultra-reliable low-latency communications* (URLLC) and mMTC. The **eMBB** service provides high-speed connectivity with increased user data rate and spectral efficiency (SE). Typical uses of eMBB are wireless broadband access and streaming

applications. On the other hand, the **URLLC** supports mission-critical applications, delivering simultaneously extremely-high reliability, on the order of $1 - 10^{-5}$, and low-latency of 1 ms or less. The associated applications include tactile internet and vehicle-to-vehicle communication to enable autonomous driving. Lastly, the **mMTC** addresses general IoT applications, which require massive connectivity and high energy efficiency, as most of the machine-type devices are powered by low-capacity batteries [7]. In the following, we explore more deeply the characteristics, requirements, technical limitations, and enabler technologies associated with the mMTC service.

1.1.1 Massive Machine-Type Communications

mMTC is a use case of the 5G wireless networks to provide communications for a massive number of devices non-controlled directly by humans. This service is designed to support the increasing demand of connecting seamlessly **huge populations of machine-type devices**, mostly related to IoT applications. These devices are associated to different fields, such as health care, home automation, manufacturing automation, and smart transportation. The main goals of mMTC is to provide communication with low-power, massive connectivity, and broad coverage. However, the traffic pattern and extremely-low complexity of the numerous machine-type devices, and the scarce radio resources result in design challenges for massive connectivity. The mMTC service aims to provide reliable and ubiquitous communication for IoT devices in the existing cellular infrastructure; however, the associated deployment issues cannot be addressed with traditional approaches.

The typical application scenario of mMTC comprises a huge population of low-cost IoT devices distributed inside a communication cell. These devices produce **sporadic and unpredictable traffic**, as they are designed to sleep most of the time to conserve battery lifetime, and activate when they have data to transmit. The machine-type devices **transmit small payloads**, are powered by low-capacity batteries, and have simple hardware. Despite these common characteristics, they have heterogeneous quality-of-service (QoS) requirements due to the different application fields for which they are intended.

The main challenges to provide mMTC are associated with the sporadic and unpredictable traffic, and the massive number of served devices. Using the conventional access strategies in this scenario with limited radio resources and the massive population implies increased scheduling delay and signaling overhead, as well as radio access network congestion. When the devices share the same pool of resources, collisions may occur as multiple users can pick the same set of resources for their access attempt. At the same time, orthogonal multiple access leads to resource underutilization due to the traffic randomness, and the inefficiency of centralized coordination and resource pre-allocation. Also, the coverage is compromised by the ubiquitous distribution of devices and their extremely low-power transmissions. It is worth mentioning that most of the results on information

theory consider small numbers of devices and asymptotically infinite-length packets, in spite of massive number of devices and short-packets, which is the operation regime of mMTC. For these reasons, the mMTC service requires dedicated theoretical foundations, resource management strategies, and physical layer technologies to achieve its goals.

The key technologies to enable mMTC implementations are associated to **multiple access schemes**. Grant-free access protocols deal with the shortage of radio resources reducing the signaling overhead. The basic strategy is to detect the set of active devices at each contention period, then decode their data. At the same time, non-orthogonal multiple access is a promising technology to increase the number of supported devices with limited radio resources. However, the high co-channel interference observed in the massive access scenario requires sophisticated interference management techniques [8]. When it comes to transceiver design, the large array gain and spatial multiplexing capability of massive MIMO can improve the coverage and increase the number of simultaneous connections. Other enablers of the mMTC service are the machine learning-assisted solutions, which have low design complexity while achieving high performance. The machine learning algorithms are applied to transceiver design, RA, channel estimation, and signal processing [8].

1.2 Extra-Large Scale Massive MIMO

The XL-MIMO is an operation regime of the massive MIMO when the physical dimensions of the array in the base-station (BS) are pushed to the extreme. This implementation integrates massive multi-antenna systems to large structures of airports, shopping malls, stadiums, and so forth to support crowded scenario locations. As the conventional massive MIMO, the XL-MIMO array provides the benefits of **large array gain**, **inter-channel orthogonality**, **channel hardening**, and low constraints related to the power amplifiers [9]. At the same time, the large array aperture provides high beamforming resolution [10], while the high number of antennas can enhance the cell coverage, improving the QoS of the border-cell users. The XL-MIMO array enables improving the QoS in crowded scenario locations, but the associated propagation channel and the high number of antenna elements imposes major implementation challenges.

The propagation channel in XL-MIMO arrays is characterized by the presence of **spatial non-stationarities**. Such non-stationarities are produced by two properties unobserved in the conventional spatial stationary massive MIMO. The first property is the operation under the **near-field propagation** regime due to the distance between the BS and the users being less than the Rayleigh distance [11]. In this case, the spherical wavefront model is suitable to represent the received signal instead of the conventional plane model. Second, each cluster of scatterers in the environment sees only a portion of the array. For this reason, the transmitted signal by each user reaches a small group of

antennas which comprises the **visibility region** of this user [12]. At the same time, the different propagation paths experienced along the array result in variations on the average received power. Results demonstrate that the non-stationarities in XL-MIMO can improve the SE if appropriated signal processing is applied, even in correlated channels [13, 14]. However, neglecting these non-stationarities can compromise seriously the performance of the system.

The implementation challenges of the XL-MIMO array are mostly associated to the high amount of data generated by the many antenna elements. The baseband data generated in the BS demands a **high interconnection bandwidth** into the data links that connect all its components. In XL-MIMO, this becomes a serious implementation bottleneck, since the required bandwidth to transmit the data from the antennas to the BS processing unit can not be handled by the current radio interfaces [15, 16]. Additionally, handling the complexity of signal processing techniques is a relevant issue, since the number of executed operations in linear detectors, such as *zero-forcing* (ZF) and *minimum mean-squared error* MMSE, scales with the number of antennas [17]. At the same time, as the conventional massive MIMO, the performance of the system is limited by the estimation of hundreds of channel links [18], a number that increases with the antennas. On the subject of hardware constraints, the fully digital implementation of XL-MIMO arrays, which require one **radio frequency (RF) transceiver** per antenna element [19], has high cost and power consumption. We also mention that there is a lack of realistic channel models for the XL-MIMO arrays that are supported by measurement campaigns.

Many transceiver architectures, algorithms, and design aspects are candidates to tackle the limitations faced by the implementation of XL-MIMO arrays. Implementing the array with limited number of RF transceivers reduces substantially the hardware cost concerns, while keeping the benefits of the high number of antennas by adopting techniques such as AS [20, 21] and hybrid precoding. The BS interconnection bandwidth problem is tackled by **distributed signal processing** and RA techniques. Adding multiple local units to pre-process the data generated by small groups of antennas reduces significantly the amount of exchanged information in the regime of asymptotic number of antennas [15, 16]. However, choosing the right signal processing techniques and coordinating adequately such local processing units is crucial to achieve this goal. At the same time, distributed signal processing is applied to perform data detection with low computational complexity by exploiting the visibility regions of the XL-MIMO arrays [22]. The literature on XL-MIMO comprises works intended to describe the XL-MIMO regime, to model the associated propagation channel, and to propose and characterize the XL-MIMO transceiver design. Also, there are works that deal with the RA problem considering AS and user scheduling.

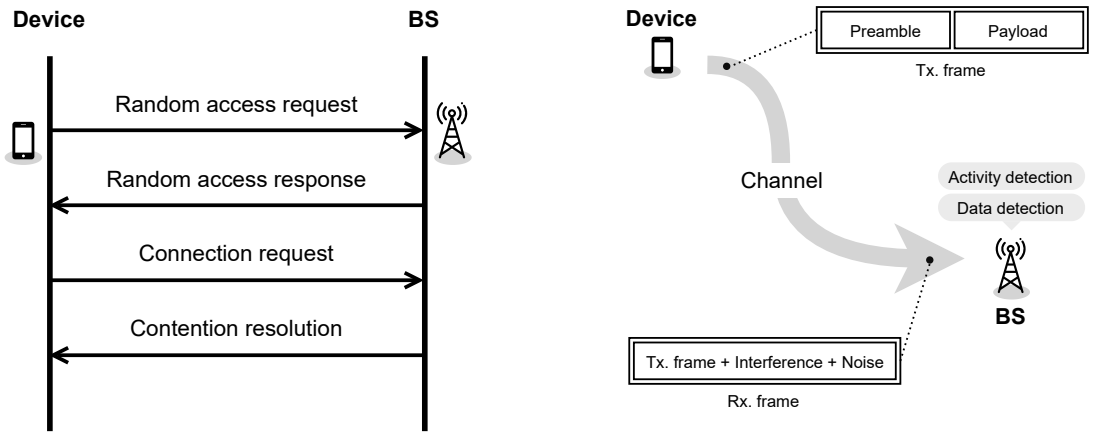
1.3 Random Access Protocols

Differently from point-to-point networks, the wireless cellular networks are cases of **multiaccess communication**, as a set of connected devices share the same communication channel. Under such scenarios, the overall performance is not limited only by the noise, distortion, and delay introduced by the channel, but also by the interfering signals transmitted simultaneously by the other nodes. Dealing with the allocation of the radio resources – *e.g.* time-frequency, pilot signals, spreading sequences – is the key path to provide efficient connectivity, being a concern of the **medium access control (MAC)** sublayer. The main problem of the multiaccess communication is to determine allowable transmission strategies for the devices capable of achieving high performance with small access delay [23].

Two classical approaches address the multiaccess communication problem. The first approach allows the devices sending their packets immediately, hoping for no interference. This strategy implies **collisions** when multiple devices use the same radio resources for transmission. Pragmatic strategies called **random access protocols** introduce mechanisms to process the collided signals and manage retransmissions in order to achieve efficient communication. However, the retransmissions must be carefully supervised, since they normally increase the access delay. The second strategy is the **channel reservation**, which aims to perfectly schedule the resources for channel use. In applications focused on massive connectivity with sporadic traffic, this strategy leads to resource underutilization [24]. In addition, considering these features with small payload sizes implies inefficiently high signaling overhead. In situations that require massive connectivity, *e.g.* IoT applications, the random access protocols are more suitable, while channel reservation becomes more adequate for cases with moderate number of devices and long-length packets, which is the case of streaming applications.

Often, the random access protocols are classified according to their grant type. In **grant-based** access protocols, an active device needs a grant response from the BS to effectively access the network. The simplest structure of a grant-based access protocol requires the exchange of four messages between the BS and an active device, as depicted in Fig. 1.1a. The grant-based protocols are known for being computationally simple due to the well-established coordination procedure. However, in massive connectivity applications, they may have increased average access delay due to the high number of collisions. At the same time, the coordination messages imply high signaling overhead. On the other hand, in **grant-free** access protocols, the active devices access the network without a grant. They send directly unique preambles for identification followed by their payload. At the BS, the active devices and their respective data are detected from the received signal and the preambles. This methodology can be deployed without coordination messages, as depicted in Fig. 1.1b. The result is the reduction on the average access delay. However, in massive connectivity situations, the number of mutually orthogonal unique preambles is

less than the number of devices due to the limited channel coherence time. For this reason, the preambles used for identification are mutually non-orthogonal, generating interference among the set of active devices. In order to tackle this scenario, the BS must implement sophisticated signal processing procedures to perform detection [8].



(a) Grant-based protocol

(b) Grant-free protocol

Figure 1.1 – Structure of two basic random access protocols according to their grant type.

1.4 Artificial Intelligence

AI is an interdisciplinary area of study dedicated to understand, model, and replicate intelligence and cognitive processes [25]. On the vision of the computer science, AI is a fundamental area for designing machines or computer programs that can compute how to act effectively in a wide variety of novel situations. These machines or programs, named **agents**, are designed with three basic abilities: **perceive the environment**, compute these perceptions in order to **adapt to changes**, and **take actions** aiming to achieve the best expected outcomes [26]. AI is used to solve different complex real-world problems, such as processing and identifying patterns in large datasets, helping the development of new drugs and the design of devices, and controlling autonomous and adaptive systems in general. On the matter of wireless communication networks, machine learning, a branch of AI, has attained the attention of researchers due to the low computational complexity of the techniques, as well as their high capability for solving intricate problems [27]. In the following, we discuss one specific class of metaheuristic evolutionary algorithms and explore the machine learning techniques, both optimization tools that will be applied in the specific XL-MIMO and mMTC random access problems in the next chapters.

1.4.1 Genetic Algorithms

GAs are a class of **metaheuristic evolutionary** search algorithms inspired by principles of genetics and natural selection often applied to solve optimization problems. They implement intelligent search techniques that explore the problem sub-spaces generating candidate solutions along the iterations, and exploit the good properties of these solutions recombining them into new ones. To operate properly, GAs do not require convex objective functions or convex constraints. Such aspect is essential as most of the practical problems on wireless communications are non-convex. Additionally, the GAs implement efficient strategies to escape from local optimums, while their overall structure being suitable for implementations which use parallel computing. Meanwhile, as other heuristic techniques, the GAs do not ensure finding the optimal solution. GAs are powerful tools to solve non-convex and combinatorial RA problems that have high dimensions with a good trade-off between performance and computational complexity.

The GAs deal with particular representations of the optimization problem and the candidate solutions. The optimization problem, including the objective functions and constraints, is cast to a fitness function which measures the quality of the solutions. At the same time, the candidate solutions are encoded into **chromosomes**, which are string representations of the optimization variables. Fig. 1.2 depicts the steps normally encountered in the standard GA. The fundamental operators of the standard GA are selection, crossover, and mutation. These operators constitute basic building blocks to implement the many variations of GAs. The **selection** operator compares the chromosomes at each iteration based on their quality (fitness), selecting the most qualified ones to generate the population of the next iteration. This operator is responsible for increasing the mean quality of solutions, as those with the best fitness are more likely to proceed as basis for the solutions of the next iteration. Common algorithmic implementations of selection are the *roulette wheel* with the fitness proportional or ranking strategies, and the tournament selection. The crossover and mutation are called variation operators, as they create diversity into the candidate solutions. In the **crossover** operator, the information contained in pairs of solutions chosen during selection are randomly merged, generating new candidate solutions. This recombination is performed intended to merge good properties of the solutions aiming to improve the overall fitness. The **mutation** operator add independent random variations on the properties of the new solutions generated during the crossover. These variations promote variability among the solutions, exploring new regions of the problem sub-spaces [28].

The GA operators introduce many input parameters to control the exploration and exploitation of the feasible sub-spaces. Hence, as other metaheuristics, the performance of a GA is strongly related to the suitable setting of these parameters. It is worth mentioning that such parameters are interrelated and they depend on the optimization problem at hand [29]. For these reasons, there is no effective rule of thumb for choosing the optimal

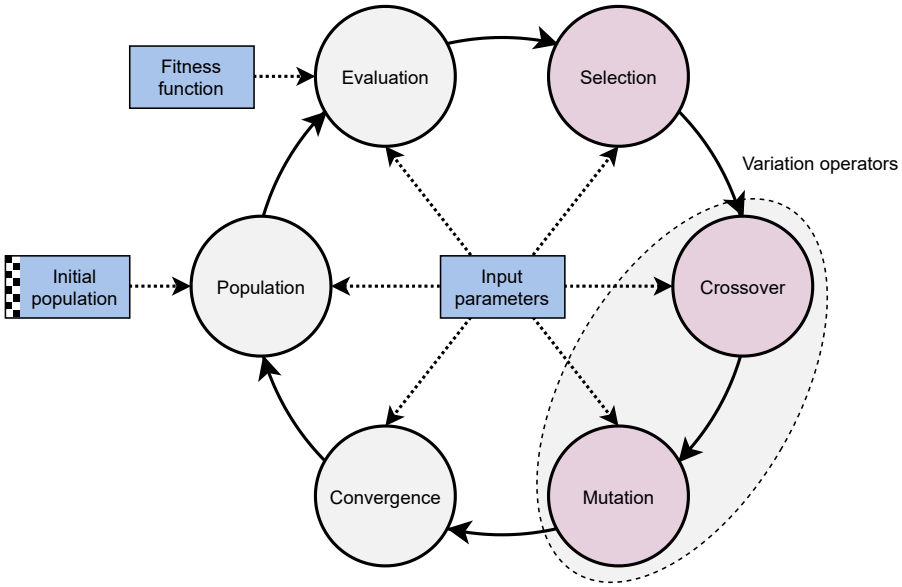


Figure 1.2 – Flowchart of the standard GA.

parameter values set. Automated tuning methods have been deployed to set the parameter values by running and evaluating different instances of a GA. Two classes of these methods are the search-based tuning and the hand-made tuning. In the search-based tuning, one heuristic search algorithm, such as the iterated local search, tries to improve the GA performance from an initial parameter configuration. In contrast, in the hand-made tuning, the designer defines a set of input parameters and iteratively modify them in order to find the combination that provides the best performance [29].

1.4.2 Machine Learning

Machine learning is a branch of AI dedicated to the study of algorithms capable of improving their performance on solving specific tasks by using a set of observations called **data set**. Most of the machine learning techniques are founded on the **generalization principle**, which consists in the ability of achieving good performance results with previously unobserved inputs. These techniques have been used extensively in applications such as computer vision, natural language processing, and anomaly detection [30]. It is worth mentioning that they have been demonstrated promising results on solving problems that have cumbersome mathematical models, as the ones observed on the study of communication networks [5]. For these reasons, machine learning algorithms are suitable candidates to be part of the future wireless networks, providing **integrated solutions** capable of optimizing simultaneously different, even conflicting, performance aspects.

The machine learning techniques are divided into three categories: supervised learning, unsupervised learning, and reinforcement learning. In the **supervised learning**, the algorithms are designed to map their inputs to specific outputs. This element becomes

clearer when we look at the typical supervised learning training set, composed by examples containing features and associated to their respective labels. This learning class is often used in classification and regression problems. In the **unsupervised learning**, the algorithms design aims to extract patterns from the data set. Differently from the first category, the unsupervised learning data sets comprise only the examples with their features [31]. Lastly, in the **reinforcement learning**, the algorithms are designed to take actions based on rewards or penalties. The basic reinforcement learning model is composed by one agent with learning capacity that interacts with the environment aiming to achieve predefined objectives. As the actions are carried out, the agent changes the environment, approaching or distancing itself from its objectives [32]. It is important to stress that reinforcement learning does not require a model to perform learning.

Despite the advancements of machine learning techniques in the computer science community, the application of these algorithms in the physical layer of wireless networks is in its initial state. Some open questions are the optimal representation of the data sets, the most suitable algorithmic architectures for each problem, and the optimal selection of loss function and training parameters, *e.g.* learning rate, training, and validation set sizes [5]. Also, it is worth mentioning that there is a lack of data sets that represent the practical operation conditions of the 5G communication networks, as well as the necessity of training methods to attain the strict real-time requirements of the communication applications [27].

1.4.2.1 Feedforward Neural Network Model

The feedforward neural network (NN) is the principal model used in machine learning for classification and regression tasks. It is employed to approximate a desired function $f^\star$ using a structure based on applying a sequence of nonlinear transformations to the input. Figuring in many commercial applications, the feedforward NN is the foundation for other specialized machine learning models, such as the convolutional NN (CNN) used widely in image processing [30]. In the following, we present the mathematical model and the main concepts associated to the feedforward NN model.

Let $\mathbf{r}_0 \in \mathbb{R}^{N_0}$ and $\mathbf{r}_{L'} \in \mathbb{R}^{N_{L'}}$ be respectively the input and output of the model. Indeed, let $f_\ell(\mathbf{r}_{\ell-1}; \theta_\ell) : \mathbb{R}^{N_{\ell-1}} \to \mathbb{R}^{N_\ell}$ be a transformation which maps the vector $\mathbf{r}_{\ell-1} \in \mathbb{R}^{N_{\ell-1}}$ to $\mathbf{r}_\ell \in \mathbb{R}^{N_\ell}$, considering θ_ℓ a set of parameters of f_ℓ. The entries of the vectors $\mathbf{r}_\ell$ are called **neurons** or **units**, while the mapping transformations in the model are called **layers**. A feedforward NN model with L' layers, namely with **depth** L', is defined by the input-output relation below:

$$\mathbf{r}_{L'} = f(\mathbf{r}_0; \boldsymbol{\theta}) = (f_1 \circ f_2 \circ \cdots \circ f_{L'-1} \circ f_{L'})(\mathbf{r}_0; \boldsymbol{\theta}) \tag{1.1}$$

where $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_{L'}\}$ is the set comprising the parameters of all the layers which compose the model. The last layer is called output layer, while the remaining $L' - 1$ intermediate

layers are known as **hidden layers**. We consider that a model with only one hidden layer is a **shallow** feedforward NN, while a model with two or more hidden layers is a **deep** feedforward NN.

The effectiveness of the feedforward NN model relies on how are defined the layer transformations $f_\ell, \forall \ell$, including the function type and the size of its output N_ℓ, namely the **width** of the layer. The format of f_ℓ strongly depends on the type of layer it implements. For instance, a densely connected layer has the format:

$$f_\ell^{\text{DC}}(\mathbf{r}_{\ell-1}; \theta_\ell^{\text{DC}}) = \sigma(\mathbf{W}_\ell \mathbf{r}_{\ell-1} + \mathbf{b}_\ell) \tag{1.2}$$

where $\sigma(\cdot)$ is the activation function, the nonlinear component of the layer, and $\mathbf{W}_\ell \in \mathbb{R}^{N_\ell \times N_{\ell-1}}$, $\mathbf{b}_\ell \in \mathbb{R}^{N_\ell}$ are respectively the network weights and the bias vector, which are the layer parameters, *i.e.* $\theta_\ell^{\text{DC}} = \{\mathbf{W}_\ell, \mathbf{b}_\ell\}$. The densely connected layer receives this name because every input neuron is connected to every output neuron. Depending on the type of tackled problem, other types of layer are used in feedforward NN models. Some examples are the convolutional layer, normalization layer, and even layers of stochastic nature as the dropout and the Gaussian noise layers.

The **activation function** contained in the layers represents the nonlinear component of the feedforward NN model. These functions are selected from a pool of effective common choices, or they are are crafted aiming to deal with specialized problems [30]. In Table 1.1 we organize the most common activation functions by their names, element-wise functions, and output ranges.

Name	**Function** $[\sigma(\mathbf{u})]_i$	**Output range**
Linear	u_i	$(-\infty, \infty)$
Rectified linear unit (ReLU)	$\max(0, u_i)$	$[0, \infty)$
Hyperbolic tangent	$\tanh(u_i)$	$(-1, 1)$
Sigmoid	$\frac{1}{1+\exp(-u_i)}$	$(0, 1)$
Softmax	$\frac{\exp(u_i)}{\sum_j \exp(u_j)}$	$(0, 1)$

Table 1.1 – List of most common activation functions [5].

Aiming to get a feedforward NN that approximates well the desired function $f^\star$, the crafted model is subjected to a **gradient-based learning** method to define the set of model parameters $\boldsymbol{\theta}$. In supervising learning applications, the learning procedure uses a **loss function** which measures the quality of the outputs predicted by the model w.r.t. the desired outputs stored in a data set. Some common choices of loss function are the cross-entropy and the mean-squared error (MSE). The set of parameters $\boldsymbol{\theta}$ is computed by minimizing the loss function with a gradient-based iterative method. Often, the stochastic gradient descent (SGD) and the adaptive moment (ADAM) estimation [33] algorithms and their variations are employed to learn the model parameters. Indeed, the **back-propagation** algorithm is employed to enable the efficient computation of the gradients in the chained structure of the feedforward NN model [30].

1.5 Research Objectives

The objectives of this Master's Dissertation are in the following:

1. **General:** Investigate the typical application scenarios of the mMTC service. At the same time, characterize the wireless networks access problem faced by this service, identifying potential grant-free random access protocols to enable connectivity. Study the XL-MIMO transceiver design as a candidate to provide connectivity in crowded communication environments. Investigate the RA problem of AS and PA in XL-MIMO arrays aiming to propose practical techniques for BS deployment;

2. **Specific:**

   (a) Perform an extensive literature review of the typical application scenarios and access protocols for the mMTC service, as well as the XL-MIMO transceiver design, its associated channel models, and signal processing algorithms;

   (b) Propose techniques to solve problems associated with the 5G wireless communications systems, specifically: **i)** grant-free random access protocols to provide massive connectivity for the mMTC service, and **ii)** RA techniques for XL-MIMO arrays to attain the bottlenecks of massive multi-antenna systems and enable practical implementations;

   (c) Implement numerical simulations in order to corroborate the effectiveness of the proposed techniques and compare them to the state-of-art in the literature. The metrics adopted for evaluation are: **i)** error probabilities during activity detection, **ii)** SE, **iii)** computational complexity in terms of floating point operations per second (flops) or execution time, and **iv)** volume of exchanged data among the BS internal components;

   (d) Disseminate the achieved results and findings on the development of grant-free random access protocols for mMTC, as well as the RA algorithms for XL-MIMO array by means of three scientific publications.

1.6 Contributions and Generated Publications

The scientific publications containing the contributions accomplished by this Master's Dissertation are:

[A] SOUZA, J. H. I. de; AMIRI, A.; ABRAO, T.; CARVALHO E. de; POPOVSKI P. Quasi-distributed antenna selection for spectral efficiency maximization in subarray switching XL-MIMO systems. *IEEE Transactions on Vehicular Technology*, v. 70, n. 7, p. 6713-6725, May 2021, DOI: 10.1109/TVT.2021.3081462 (IF: 5.379, Eng. IV Qualis-CAPES: A1).

[B] SOUZA, J. H. I. de; ABRAO, T. Deep learning-based activity detection for grant-free random access. *Submitted in 05/17/2021 to the IEEE Transactions on Vehicular Technology* (IF: 5.379, Eng. IV Qualis-CAPES: A1).

[C] SOUZA, J. H. I. de; ABRAO, T. Performance of Activity Detection and Channel Estimation Methods for Machine-Type Communications with MIMO Transceiver. *Manuscript in preparation for submission.*

1.7 Organization of the Text

The remainder of this text is organized as follows. In Chapters 2, 3, and 4 we outline the results obtained by the publications, emphasizing their contributions. In Chapter 5 we provide the final comments, as well as devise possible paths for future works involving the discussed topics. Lastly, in Appendices A, B, and C are reproduced, in chronological order of conception, the full text of the scientific publications associated to this Master's Dissertation.

2 Antenna Selection for Spectral Efficiency Maximization in XL-MIMO Systems

In the work in Appendix A, we study the joint AS and PA problem for an XL-MIMO system with limited number of RF transceivers at the BS. The aim of the addressed RA problem is to jointly select the set of active antennas and distribute the available power among them in order to maximize the SE of the system, given the constraints of maximum number of RF transceivers and limited connections imposed by subarray structure in the BS. The contributions of the work are as follows:

(a) Description of a distributed transceiver design for XL-MIMO based on a subarray switching architecture, as depicted in the right side of Fig. 2.1;

(b) Proposition of a centralized procedure based on the evolutionary metaheuristic GA to perform joint AS and PA to maximize the SE with subarray connection and maximum transmitted power constraints;

(c) Proposition of a distributed version of the GA procedure for joint AS and PA which achieves performance tight to the centralized one, but with low-size coordination data and less number of executed operations;

(d) Extensive analysis of the proposed procedures in terms of number of symbols for training, coordination data size, and number of flops.

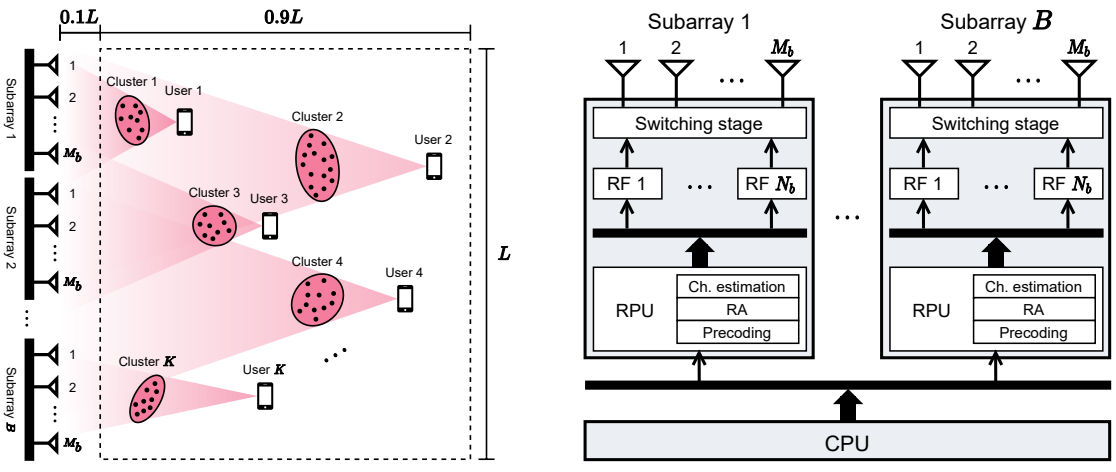


Figure 2.1 – Diagrams of the XL-MIMO communication cell (left) and the BS based on a subarray switching structure during the downlink (right).

The studied XL-MIMO system is depicted in the left side of Fig. 2.1, while the BS structure during the downlink is illustrated in the right side. The BS is composed by B independent subarrays, each one equipped with a remote processing unit (RPU), N_b RF transceivers, and $M_b > N_b$ antennas. The RPUs locally carry out channel estimation, RA, and precoding calculation. In addition, a flexible switching stage connects each active antenna of the subarray to one available RF transceiver. A central processing unit (CPU) coordinates the operation of the multiple subarrays in the BS.

Since the number of RF transceivers is less than the number of antennas, AS is crucial to attain a reasonable downlink performance. On this matter, we devise two GAs to perform joint AS and PA to maximize the SE, given the maximum number of RF transceivers, the subarray connection constraints, and the limited power for transmission. The GAs are designed to compute near-optimal solutions of the optimization problem in eq. (15, Appendix A), reproduced in the following for the sake of clarity.

$$\underset{\mathbf{D},\mathbf{P}}{\text{maximize}} \ \text{SE} = \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k}{\sigma_z^2}\right) \tag{2.1a}$$

$$\text{subject to} \ \sum_{m \in \mathcal{M}_b} D_m \leq N_b, \qquad b = 1, \ldots, B \tag{2.1b}$$

$$\text{tr}\left[\mathbf{P}(\mathbf{H}^H \mathbf{D}\mathbf{H})^{-1}\right] \leq P_{\max} \tag{2.1c}$$

$$D_m \in \{0, 1\}, \qquad m = 1, \ldots, M \tag{2.1d}$$

$$p_k \geq 0, \qquad k = 1, \ldots, K \tag{2.1e}$$

p_k is the allocated power for users k, while σ_z^2 is the thermal noise power. $D_m, \forall m$ are the state indicators of the antennas which assume value equal to 1 if the antenna is active, or equal to 0 otherwise. $P_{\max}$ is the available power for transmission, while $\mathbf{H}$, $\mathbf{P}$ and $\mathbf{D}$ are, respectively, the channel matrix, the diagonal matrix with the elements $p_k, \forall k$, and the diagonal matrix with the elements $D_m, \forall m$. In eq. (2.1), the objective function is the SE obtained by the ZF precoding. At the same time, the constraints are, respectively, the connection constraints associated to the switching stage of each subarray, the maximum power constraint, and the restriction of the antenna state indicators to binary values.

The *GA for RA* (GA-RA) is the first proposed algorithm. Executed entirely in the CPU, it uses full knowledge of the channel state information (CSI) to compute the indices of the active antennas and the power distribution. Aware of the interconnection bandwidth problem associated to obtaining full CSI at the CPU in systems with extremely large number of antennas, we propose a distributed GA called *quasi-distributed GA for RA* (DGA-RA). The core idea of the DGA-RA is to independently carry out AS and PA at each subarray, using only local CSI and side information provided by the CPU. The local RA is enabled by the Sherman-Morrison-Woodbury formula, a matrix identity that allows calculating $(\mathbf{R} + \mathbf{p}\mathbf{q}^H)^{-1}$ from $\mathbf{R}^{-1}$, $\mathbf{p}$ and $\mathbf{q}$. Fig. 2.2 depicts the iterative procedure of the DGA-RA algorithm. Details on the computations are shown in Section III-B of Appendix A.

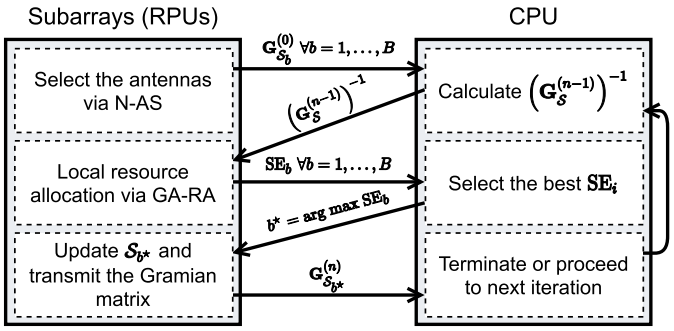
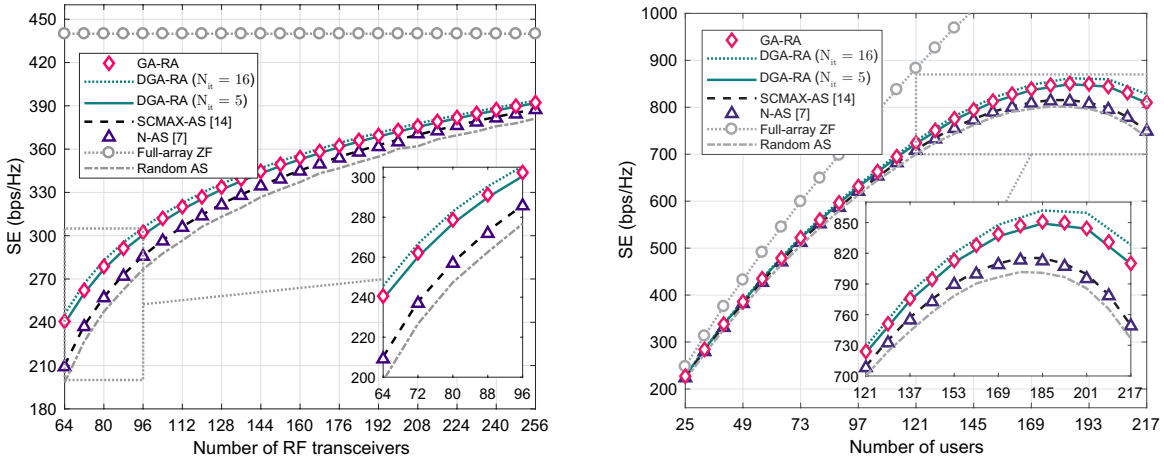


Figure 2.2 – Diagram of the proposed DGA-RA procedure steps with coordination between the CPU and the RPUs.

The proposed techniques are evaluated according to their performance and complexity. Also, we compare them with two methods of the literature. One of the baseline methods is a heuristic based on the power of the channel coefficients, called *norm AS* (N-AS) [19]. The other is based on the convex optimization framework, intended to maximize the downlink sum-capacity rather than the SE [34]. We call it the *sum-capacity maximization AS* (SCMAX-AS) method.

Fig. 2.3a depicts the SE *versus* the number of RF transceivers, while the Fig. 2.3b depicts the SE *versus* the number of users. The results for DGA-RA are generated considering 5 and 16 iterations of the procedure illustrated in Fig. 2.2. Besides, it is worth mentioning that we plot the results for the full-array ZF and random AS as upper and lower performance bounds, respectively. Analyzing the Fig. 2.3a, we see that the GA-based techniques outperform the baseline methods. The performance gap reduces as the number of RF transceivers increase. Comparing the GA-based techniques, we see that the DGA-RA with 5 iterations almost equals the GA-RA, while its version with 16 iterations outperforms the results of the GA-RA. Next, in Fig, 2.3b we perceive that the performance increases with the number of users, up to a point where such behavior reverses. This is due to the reduction of the spatial degrees of freedom of the system by increasing the number of users while keeping the number of activated antennas constant. When the number of users is high, *i.e.* in crowded scenarios, the GA-based techniques operate substantially better than the baseline methods.

Lastly, the Fig. 2.4 addresses the results on complexity. Fig. 2.4a depicts the coordination data size *versus* the number of users and number of antennas, considering different number of iterations of the DGA-RA algorithm, antennas, and subarrays. We consider the coordination data size as the amount of samples transmitted to the CPU during the RA. This result evidences that the coordination data size of the DGA-RA technique does not depend on the number of antennas, and is lower than the centralized methods when the number of users is small. Fig. 2.4b depicts the number of flops per processing unit *versus* the number of RF transceivers and users for the GA-RA and DGA-RA techniques. Analyzing this result, we see that, when the number of RF transceivers

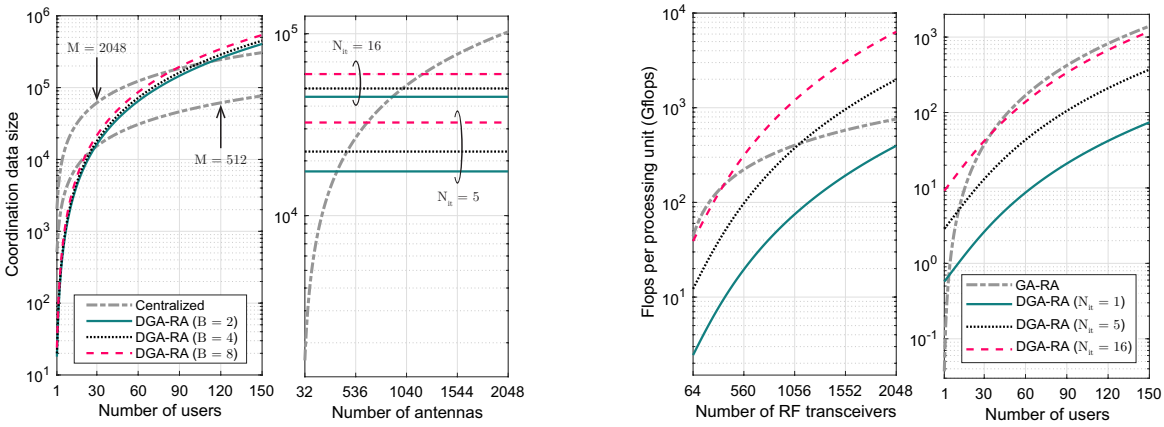


(a) SE *vs.* the number of RF transceivers

(b) SE *vs.* the number of users

Figure 2.3 – Performance of the proposed DGA-RA, GA-RA algorithms, and the N-AS and SCMAX-AS baseline techniques for AS to maximize the SE in XL-MIMO systems. The SE attained by full-array ZF and random AS are plotted as a lower and upper performance bounds, respectively. The parameters values aret $M = 512$, $B = 8$ and, when it is not specified, $K = 50$ and $N = 256$.

is low and the number of users is high, the DGA-RA technique has advantage over the GA-RA.



(a) Coordination data size *vs.* the number of users and the number of antennas

(b) Flops per processing unit *vs.* the number of RF transceivers and the number of users

Figure 2.4 – Complexity of the proposed DGA-RA, GA-RA algorithms, and the baseline techniques schemes for AS to maximize the SE in XL-MIMO systems. In (a), when it is not specified, $N_{it} = 16$ and $K = 50$, while the centralized techniques include the GA-RA, the SCMAX-AS and the full-array ZF. In (b), $B = 8$ and, when it is not specified, $K = 50$ and $N = 256$.

2.1 Conclusions

In the work in Appendix A, we propose a subarray switching architecture for the BS antenna array and investigate the joint problem of AS and PA in XL-MIMO systems with limited number of RF transceivers. We outline the main findings and conclusions of the publication in the following:

- Two GA-based procedures that achieve near-optimal performance with low computational complexity are proposed, one implemented with a centralized approach and the other with a distributed scheme;
- Numerical simulations corroborate that the proposed GA-based procedures attain high performance-complexity trade-off when compared with the N-AS and SCMAX-AS baseline techniques, specially in crowded XL-MIMO scenarios;
- The DGA-RA fulfills the initial proposal to perform AS and PA transmitting a reduced amount of samples to the CPU. Additionally, we demonstrate that the distributed procedure outperforms the GA-RA in both performance and complexity by taking the appropriate system operation settings.

3 Deep Learning-based Activity Detection for mMTC Grant-free Random Access

In the work in Appendix B, we study the problem of device activity detection under grant-free random access protocols aiming to support mMTC scenarios. The contributions of the work are as follows:

(a) Proposition of two data-driven deep learning (DL) algorithms for activity detection of machine-type devices, one based on the densely connected layers and the other based on convolutional layers;

(b) Extensive description of the proposed NNs, presenting the procedures for hyper-parameter tuning and network training;

(c) Carrying out experiments to evaluate the performance of the activity detection algorithms in terms of error rate and computational complexity, and compare to techniques in the recent literature;

(d) Evaluation of the proposed DL algorithms for activity detection with three kinds of preamble sequences: Gaussian and Bernoulli random sequences, as well as deterministic Zadoff-Chu sequences.

The addressed communication scenario is illustrated in Fig. 3.1, including the random access slot and a high-level diagram of the DL-based activity detection algorithm. During the uplink, each active device uses the random access slot to transmit its unique preamble with a length of L symbols followed by the payload. The received preambles are used to carry out activity detection and channel estimation aiming to gather sufficient information to decode the data transmitted by each active device. We focus on developing DL-based techniques to perform activity detection considering mMTC communications scenarios.

The design of the preamble sequences is non-orthogonal, since the size of the population of devices, K, is greater or even much higher (crowded scenarios) than the number of symbols of the preamble sequences, L. This assumption is consistent with practical applications as one of the mMTC application aims to handle huge populations of machine-type devices. However, the non-orthogonality makes the activity detection challenging, since the multiple active users suffer from high inter-user interference levels.

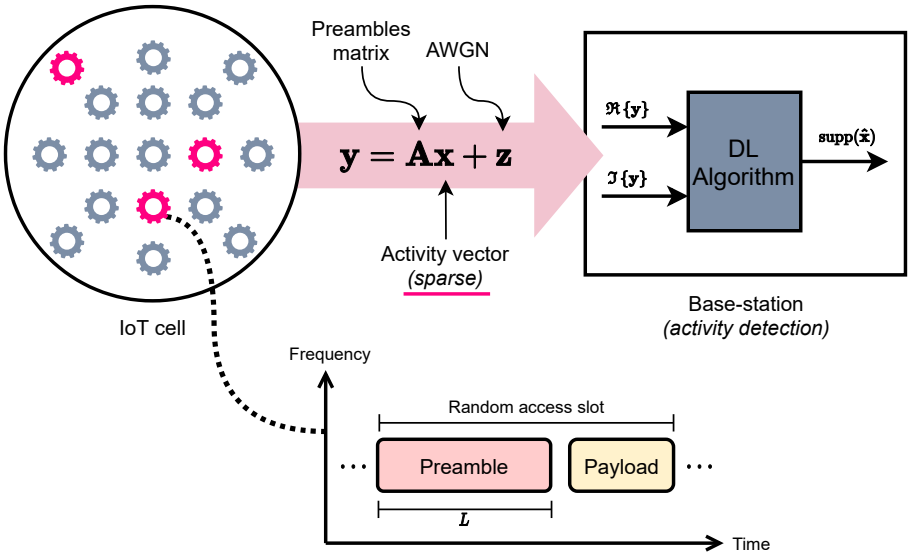


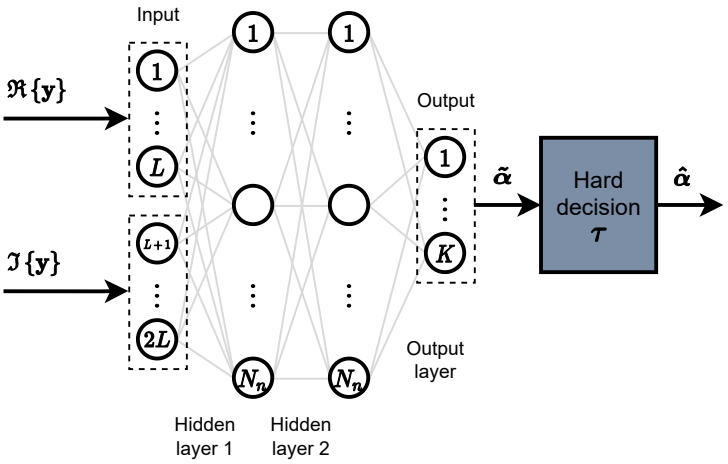
Figure 3.1 – Diagram of the communication scenario with the random access slot and an activity detection scheme based on a DL algorithm.

The proposed schemes for activity detection are a *DFN* and a *CNN*, depicted respectively in Figs. 3.2a and 3.2b. Such NNs have extremely-low run time at the cost of an intense off-line training phase to learn their parameters.

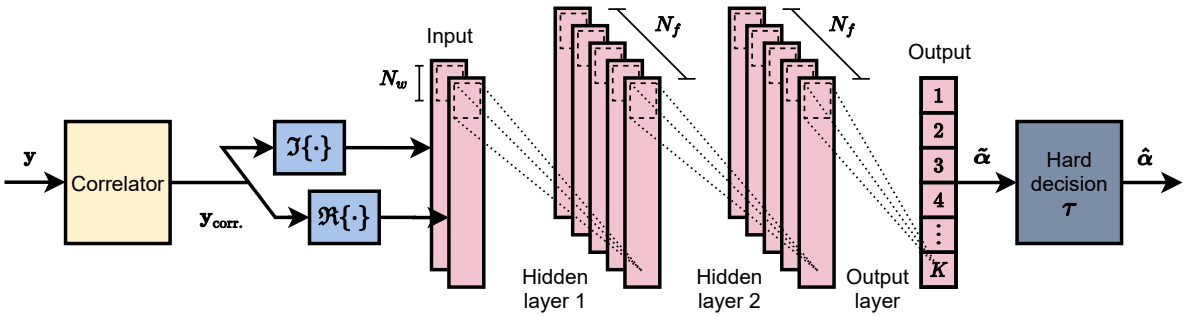
The inputs of the DFN are the real and imaginary parts of the received signal. The network has 3 densely connected layers with N_n, N_n, and K neurons, respectively. Besides, the hidden layers have the ReLU activation function, while the output layer have the sigmoid activation function. On the other hand, the input of the CNN is a correlator stage, that multiplies the received signal by the conjugate transpose of the preamble matrix. The network is composed by the 3 layers of the type 1D convolution. The hidden layers have N_f feature maps and ReLU activation function. At the same time, the output layer has 1 feature map and sigmoid activation function. The filter length of all the layers is equal to N_w. Also, both DFN and CNN have a hard decision module at their outputs with parameter $\tau > 0$. We incorporate this module since the original activity descriptors are binary, while the outputs of the networks are values in the continuous range $(0, 1)$.

The proposed DL-based techniques are evaluated and compared with two methods recently available in the literature, the *least absolute shrinkage and selection operator* (LASSO) [35] and the *approximate message passing* (AMP) algorithm [36]. Additionally, we study the performance of the proposed techniques with three types of preamble sequences: random sequences generated by a symmetric Bernoulli distribution, by a complex Normal distribution, and deterministic Zadoff-Chu sequences [37].

Fig. 3.3 depicts the detection error rate of the evaluated techniques *versus* the undersampling ratio Δ, calculated by dividing the preamble length by the total number of devices. We see on this results that the DL-based techniques achieve competitive performance when compared to the state-of-art AMP and the LASSO methods. Comparing



(a) DFN



(b) CNN

Figure 3.2 – Diagrams of the proposed deep NNs for activity detection of machine-type devices.

the proposed methods, we perceive that the DFN with the three types of sequences achieves performance substantially higher than the CNN with random sequences. However, the CNN with the Zadoff-Chu sequences attains a significant improvement on performance, reaching error rate values comparable to the DFN. Such fact is due to the correlator stage at the input of the CNN. Although the LASSO attains the best performance for $\Delta > 0.175$, we demonstrate in the subsequent results that this technique has extremely high computational complexity, resulting in a poor performance-complexity trade-off.

Next, Fig. 3.4 depicts the receiver operating characteristic (ROC) curves of the evaluated techniques for two values of undersampling ratio. For $\Delta = 0.525$, the CNN with Zadoff-Chu sequences has a huge improvement on performance, if compared with the random sequences. Additionally, the result of the CNN is tight to the AMP algorithm. At the same time, the DFN attains a performance superior to the AMP and comparable with LASSO. On the other hand, for $\Delta = 0.175$ all the evaluated techniques have comparable performance, except the AMP algorithm that attains significantly degraded results.

In Fig. 3.5 are depicted the detection error trade-off (DET) curves of the evaluated

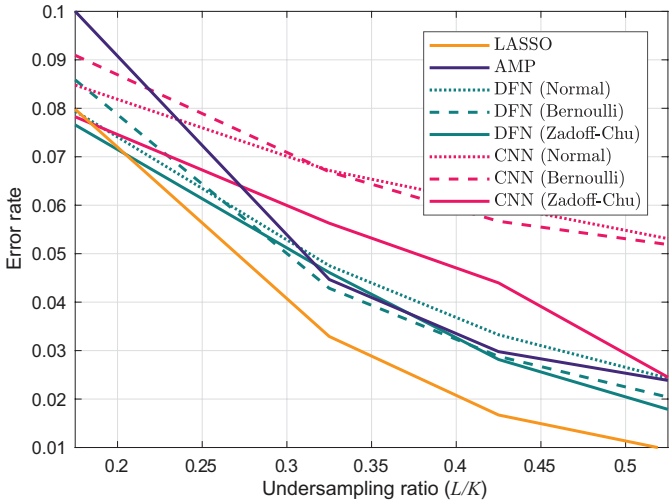
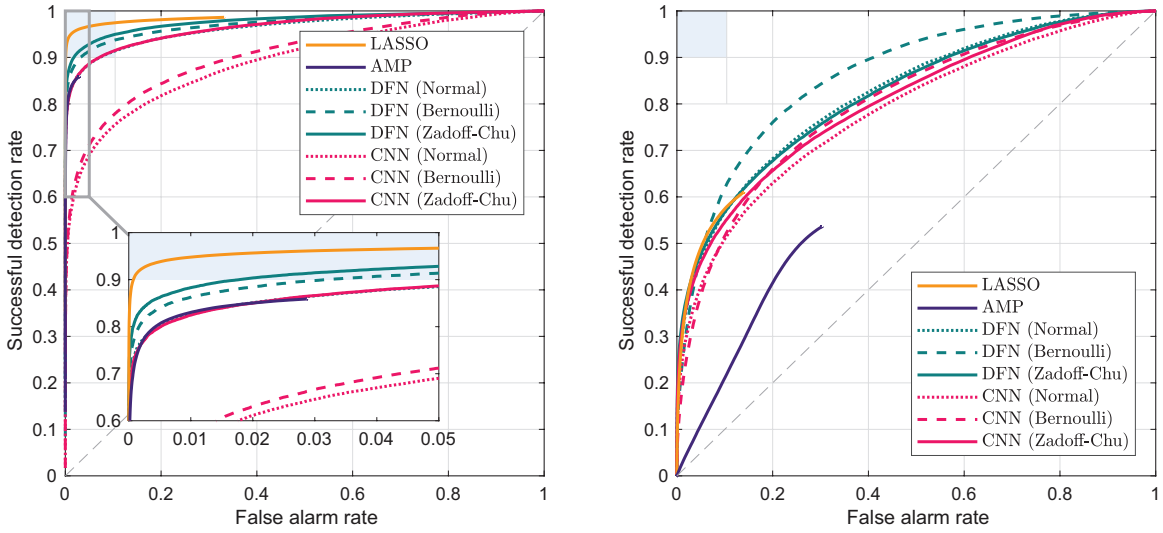


Figure 3.3 – Detection error rate *vs.* the preamble length of the proposed DL-based algorithms using the Normal, Bernoulli, and Zadoff-Chu sequences, and of the LASSO and AMP baseline techniques.

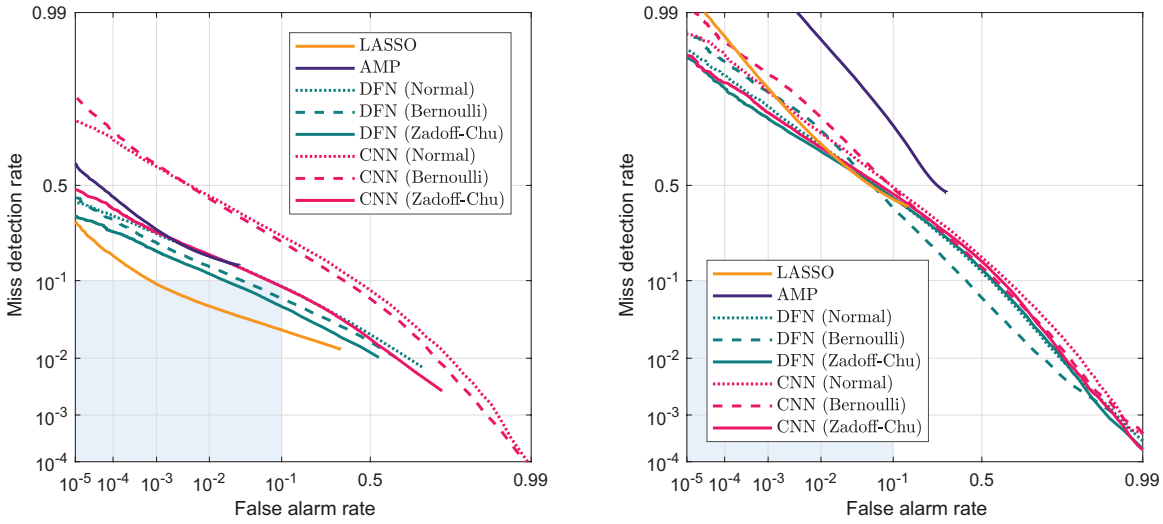


(a) $\Delta = 0.525$ (b) $\Delta = 0.175$

Figure 3.4 – ROC curves of the proposed DL-based algorithms using the Normal, Bernoulli, and Zadoff-Chu sequences, and of the LASSO and AMP baseline techniques.

techniques for two values of undersampling ratio. Such result is similar to the one illustrated in Fig. 3.4, with the difference that the axes represent the two types of error that occur during activity detection. We see that the proposed DL-based techniques cover wider ranges of miss detection and false alarm rates than the baseline techniques. Additionally, for $\Delta = 0.525$ all the techniques achieve smaller false alarm rates, owing to the sporadic activity of the devices.

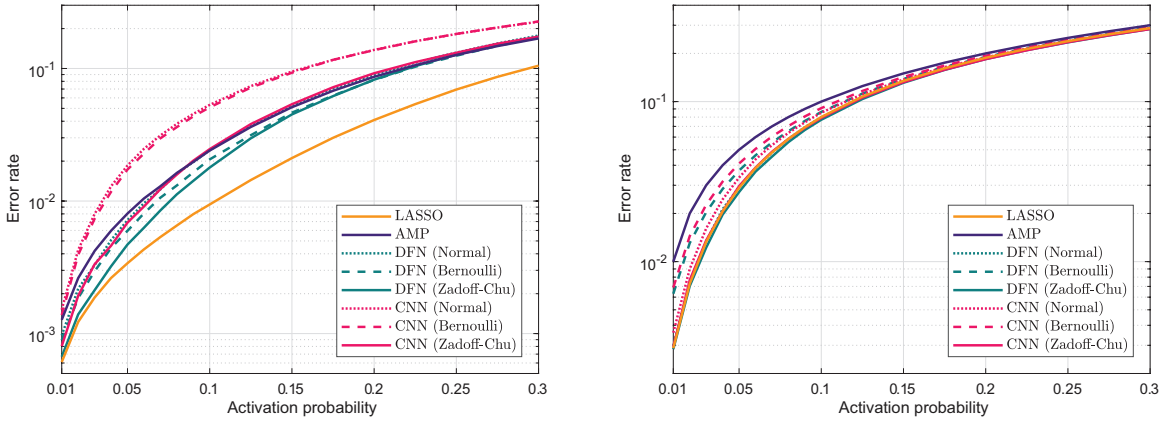
Fig. 3.6 depicts the detection error rate *versus* the activation probability of the evaluated techniques for two values of undersampling ratio. We stress that the DL-based techniques are trained with samples with a constant activation probability equal to 0.1. Hence, this result demonstrates the robustness of the DL-based activity detection tech-



(a) $\Delta = 0.525$ (b) $\Delta = 0.175$

Figure 3.5 – DET curves of both proposed DL-based algorithms using the Normal, Bernoulli, and Zadoff-Chu sequences, and of the LASSO and AMP baseline techniques.

niques against variations on the operation scenario. The proposed techniques demonstrate consistent performance, following the LASSO and AMP behavior as the activation probability increases. For $\Delta = 0.525$, we see again the improvement of the CNN using Zadoff-Chu sequences. On the other hand, for $\Delta = 0.175$, the performance of all the techniques is similar for activation probability values greater than 0.1.

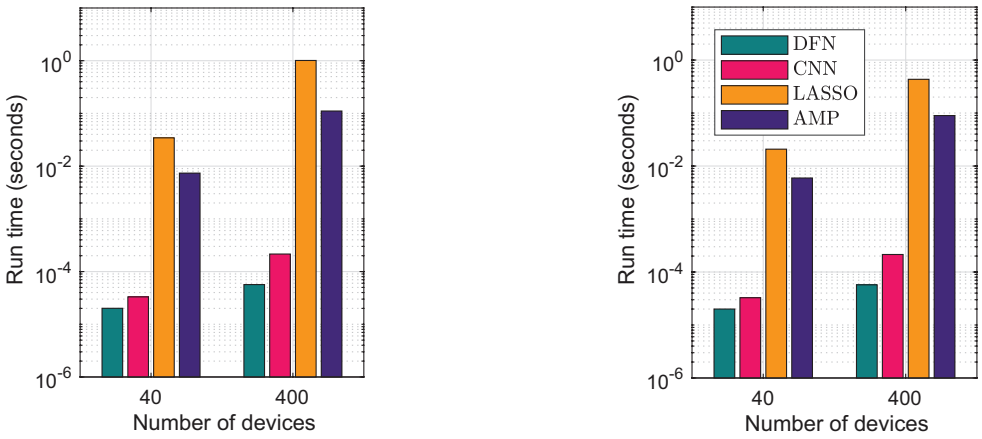


(a) $\Delta = 0.525$ (b) $\Delta = 0.175$

Figure 3.6 – Detection error rate *vs.* the activation probability of the proposed DL-based activity detection algorithms using the Normal, Bernoulli, and Zadoff-Chu sequences, and of the LASSO and AMP the baseline techniques.

On the matter of computational complexity, Fig. 3.7 illustrates the run time *versus* the number of devices of the evaluated techniques for two values of undersampling ratio. The proposed techniques have extremely low run time values, on the order of 10 μs to 100 μs, at least two orders of magnitude less if compared with both baseline methods.

Besides, the CNN have complexity slightly higher than the DFN. We notice that the complexity does not change drastically with the undersampling ratio, but increases with the number of devices, as the number of computed outputs grows. Finally, we notice that, when compared to the other evaluated techniques, the good performance of the LASSO method comes with an extremely high computational complexity.



(a) $\Delta \approx 0.17$ (b) $\Delta \approx 0.52$

Figure 3.7 – Run time *vs.* the number of devices of the proposed DL-based algorithms, and of the LASSO and AMP baseline techniques, considering two values for the undersampling ratio.

3.1 Conclusions

In the publication available in Appendix B, the problem of device activity detection in the context of grant-free random access for the mMTC scenarios is investigated. We propose two DL-based procedures to perform activity detection. The main findings and conclusions of the work are outlined in the following.

- A DFN based on densely connected layers and a CNN based on 1D convolutional layers are proposed for activity detection of machine-type devices;
- The impact of the type of the preamble sequences on the performance of the proposed algorithms is analyzed. We investigate the performance of the DFN and CNN with random sequences generated by sampling a complex Normal and a symmetric Bernoulli distributions, and deterministic Zadoff-Chu sequences;
- Numerical experiments reveal that the DL-based algorithms attains extremely-low computational complexity when compared with the state-of-art LASSO and AMP methods. Additionally, the DFN achieves the best performance while the CNN with Zadoff-Chu sequences outperforms substantially the CNN using random preambles.

4 Activity Detection and Channel Estimation for mMTC Grant-free Random Access

In the work in Appendix C, we investigate the problem of device activity detection and channel estimation for grant-free random access in mMTC applications. The contributions of the work are as follows:

(a) Revision and introduction of four methods for activity detection and channel estimation in the context of grant-free random access;

(b) Systematic evaluation and comparison of the introduced techniques in order to analyze their respective performance and main operation regimes.

We consider the uplink transmission of a cellular mMTC system served by a BS equipped with M antennas. During the random access slots, the active devices follow a grant-free random access scheme, transmitting unique mutually non-orthogonal pilot sequences for activity detection and channel estimation. Fig. 4.1 depicts a diagram of the investigated communication scenario during the described pilot phase. The detailed definition of the parameters of the communication scenario is available in Section II of Appendix C.

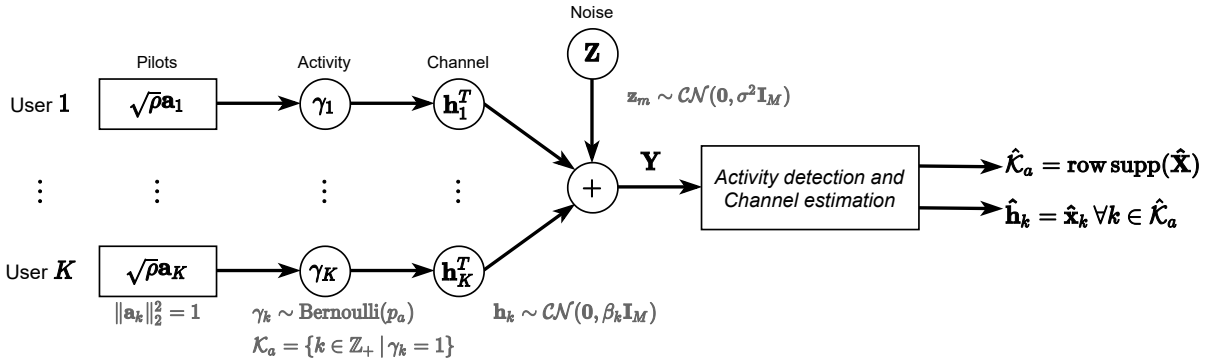


Figure 4.1 – Diagram of the pilot phase, in which the BS perform activity detection and channel estimation. The output of the activity detection and channel estimation module is a set with the indices of the active users and their respective estimated channel vectors.

Now, we briefly describe the evaluated techniques. The *MMSE* estimator is a classical estimation method which can be used to estimate the channel vectors after the activity detection. The derivation of such method is done by calculating the closed-form solution of the optimization problem of minimizing the expected value of the MSE

between the received signal and its reconstruction from the estimated channel vectors. It is worth mentioning that the MMSE estimation requires knowledge of the channel statistics. Differently, the *group LASSO* method [38] consists in the solution of an optimization problem with an objective function that can be divided into a sum of two parts. The first part represents the MSE between the received signal and its reconstruction from the estimated channel vectors. The second part is a scaled sum of the norms of the estimated channel vectors. Specifically, the second part promotes the sparsity of the solution. This is favorable for our estimation problem, since the desired signal has a sparse structure. The covariance-based estimation method consists in the optimization problem of minimizing the MSE between the expected value of the covariance matrix of the received signal and an estimate of it. Such problem formulation is equivalent to the *non-negative least-squares* (NNLS) optimization program [39], and can be solved by coordinate descent (CD) methods [40]. Finally, the *vector AMP* algorithm [41] is a multiple-measurement vector (MMV) version of the original AMP [36]. Such technique is specialized on the estimation of sparse signals by using an iterative procedure. The state evolution analysis in the AMP theory allows theoretically predicting the performance of the technique [36].

The introduced techniques are evaluated and compared in terms of the normalized MSE (NMSE) of the estimated channel vectors and the detection error rate. On the matter of the error rate, both ROC and DET curves are constructed. For a complete description of the simulation scenario, refer to the Section IV-C and the Table II in Appendix C. Fig. 4.2 depicts the NMSE *versus* the undersampling ratio of the genie-aided MMSE and the group LASSO. Such result is generated for two values of antennas at the BS. Both techniques reach an NMSE floor after a certain value of undersampling ratio. Therefore, the performance of long pilot sequences is limited. The genie-aided MMSE, which assumes perfect knowledge of the users activity and the channel statistics, achieves the best performance. However, we notice that the number of antennas does not have impacts on the NMSE for such technique. Conversely, the group LASSO improves the performance with 64 antennas and $\Delta \geq 0.2$. The exception is the point where $\Delta = 0.1$, in which the group LASSO with 4 antennas outperforms the one with 64.

Next, Fig. 4.3 depicts the detection error rate *versus* the undersampling ratio of the NNLS and the group LASSO, considering two values for the number of antennas at the BS. Again, the performance increases with the undersampling ratio, but it is limited. The group LASSO performance is considerably higher than the NNLS for $\Delta \geq 0.2$. On the matter of the number of antennas, the NNLS with 64 antennas achieves better results than with 4 antennas. The same occur with the group LASSO for $\Delta \geq 0.3$. Such behavior reverses for $\Delta < 0.3$.

Still on the performance of activity detection, Fig. 4.4 depicts the ROC curves of the NNLS and group LASSO techniques, considering two values of antennas at the BS and two values of undersampling ratio. The result demonstrates that, in general, under

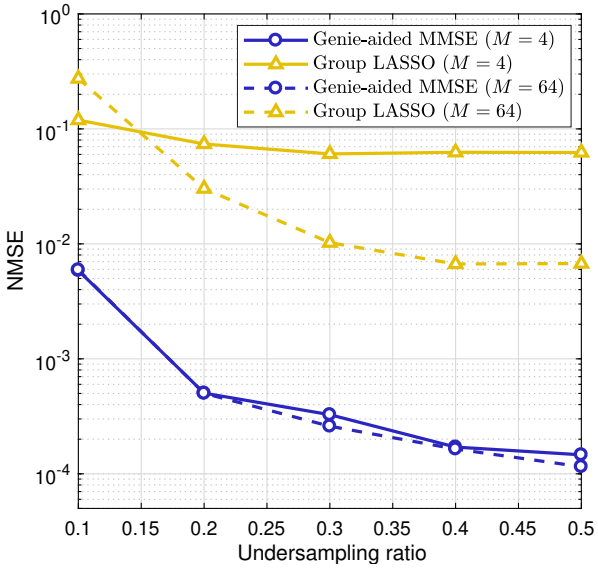


Figure 4.2 – NMSE of the estimated channel vectors of the detected users *vs.* the undersampling ratio (L/K) of the genie-aided MMSE and the group LASSO. Two values for the number of antennas at the BS are considered.

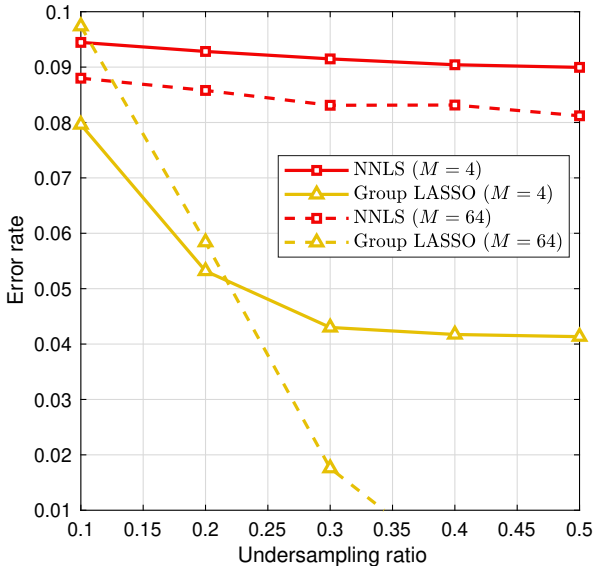
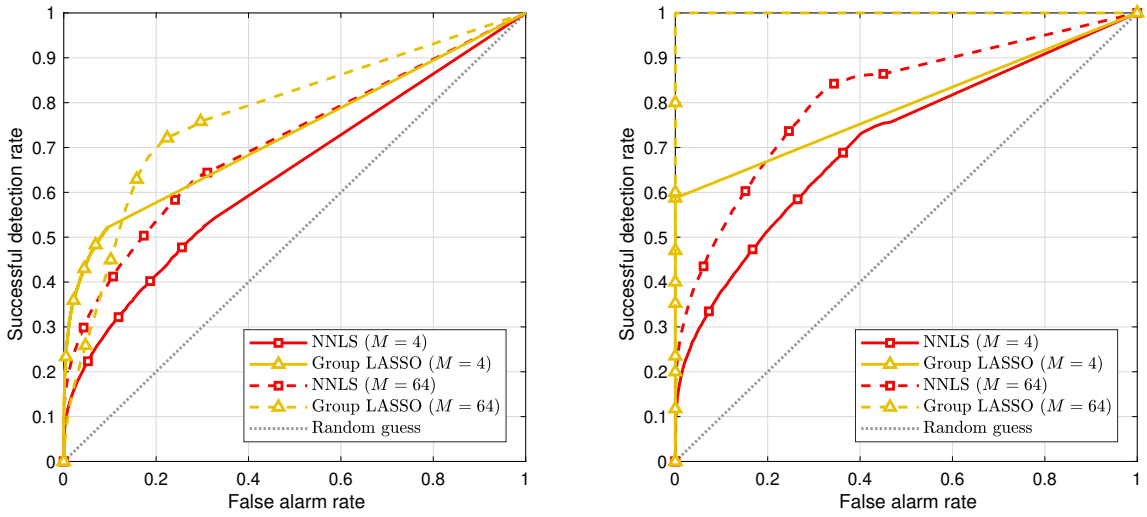


Figure 4.3 – Detection error rate *vs.* the undersampling ratio (L/K) of the NNLS and group LASSO considering two values of antennas at the BS.

the same conditions, the group LASSO attains better ROC curves than the NNLS. All the curves improve by increasing the number of antennas, except for the group LASSO for $\Delta = 0.1$. In the particular case of group LASSO for $\Delta = 0.1$, 4 antennas provide low false alarm rates, while 64 antennas provide high successful detection rates.

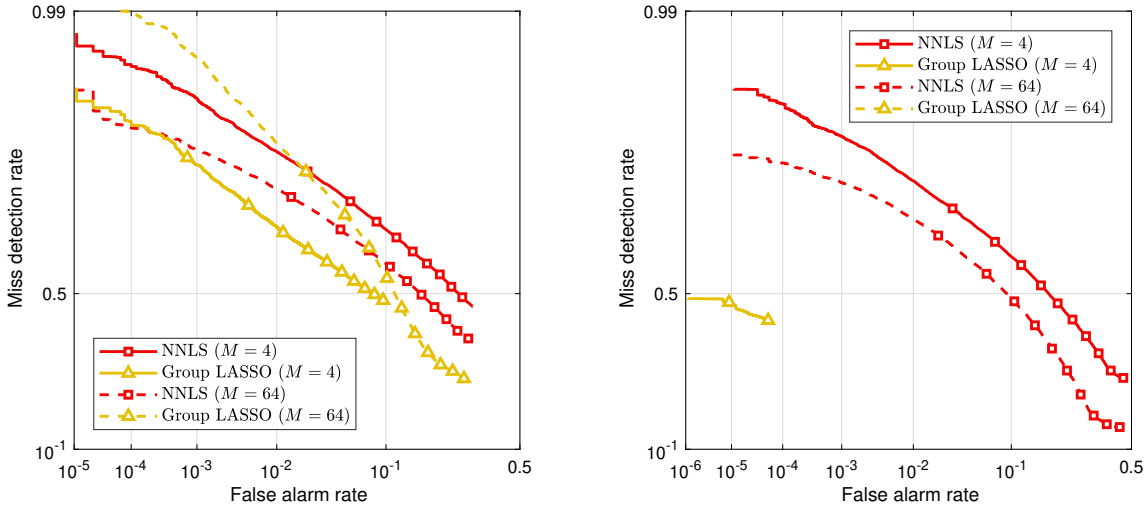
Fig. 4.5 depicts the DET curves of the NNLS and group LASSO techniques, considering two values of antennas at the BS and two values of undersampling ratio. Analyzing the results, we notice that, except for the group LASSO for $\Delta = 0.1$, both false alarm and miss detection rates improve by increasing the number of antennas. In the



(a) $\Delta = 0.1$ (b) $\Delta = 0.5$

Figure 4.4 – ROC curves of the NNLS and group LASSO for two values of antennas at the BS and considering two values of undersampling ratio (L/K). The dotted curves represent the performance of choosing the activity descriptors randomly.

exceptional case, the group LASSO degrades with 64 antennas when $\Delta = 0.1$.



(a) $\Delta = 0.1$ (b) $\Delta = 0.5$

Figure 4.5 – DET curves of the NNLS and group LASSO for two values of antennas at the BS and considering two values of undersampling ratio (L/K).

4.1 Conclusions

In the publication in Appendix C, we investigate the problem of activity detection and channel estimation in grant-free random access protocols for mMTC applications. The evaluated techniques are the MMSE estimation, the group LASSO method, the covariance

matrix estimation formulated as a NNLS optimization problem, and the VAMP algorithm. The main findings and conclusions of the work are outlined in the following:

- The four techniques are introduced and compared in terms of the NMSE of the estimated channel vectors and the device detection error rate;
- The numerical experiments demonstrate that the group LASSO technique with sufficiently long pilots benefits consistently from increasing the number of antennas. Still with long pilots, the group LASSO technique achieves good results on activity detection. Lastly, the NNLS technique improves the activity detection performance by increasing the number of antennas.

5 Conclusions

In this Master's Dissertation, we investigate the typical use case of the mMTC service, analyzing deeply different types of grant-free random access protocols to handle large groups of machine-type devices. On this matter, we propose grant-free random access protocols based on machine learning. Also, we investigate the XL-MIMO transceiver, which is a strong candidate to handle crowded communication scenarios. We study the RA problem of joint AS and PA in XL-MIMO arrays, proposing distributed and centralized solutions based on the GAs. The results of this study are presented by means of three scientific publications addressing the listed issues, proposing techniques to enable the massive connectivity in wireless communication networks. The main conclusions of these three works are:

- The XL-MIMO BS with a limited number of RF transceivers benefits efficiently from the extremely-high number of antennas with GA-based AS procedures. The proposed GAs achieve near-optimal performance with low computational complexity. Besides, the distributed GA fulfills the technical limitations of the XL-MIMO, reducing the size of the coordination data which need to be transferred to the CPU;

- DL-based solutions demonstrate a good performance-complexity trade-off on activity detection for grant-free random access, attaining the performance of state-of-art techniques with extremely-low computational complexity. Additionally, we notice the impact of the type of the preamble sequences on the performance of the DL-base technique, as the Zadoff-Chu sequences, which present good cross-correlation properties, outperform significantly the performance achieved by random sequences;

- On the matter of device activity detection and channel estimation for grant-free random access, techniques available in the recent literature benefit from the multiple antenna transceiver, improving their performance even using pilot sequences with limited length.

5.1 Future Research Directions

Finally, we highlight some future research directions associated with the study accomplished in this Master's Dissertation:

- **DL-based activity detection with MIMO:** few works approach the activity detection problem with machine learning techniques considering the MIMO transceiver. Furthermore, the existing works approach the deep autoencoder model [42], or

model-driven methods obtained by unfolding iterative algorithms [43]. Other machine learning models need to be evaluated. For instance, models developed with convolutional layers are suitable to process the uplink signal simultaneously in the time and antenna domains;

- **DL-based random access protocols for heterogeneous applications:** IoT exhibits a diversity of supported devices and applications, with varying QoS requirements, demands, and traffic models. Solutions tackling the problems associated to such heterogeneous requirements are needed [44]. DL models are promising for implementing efficient resource – *e.g.* pilots, and time-frequency slots – allocation strategies, and receiver designs accounting the different activation patterns and priorities of devices.

Bibliography

1 OSSEIRAN, A.; BOCCARDI, F.; BRAUN, V.; KUSUME, K.; MARSCH, P.; MATERNIA, M.; QUESETH, O.; SCHELLMANN, M.; SCHOTTEN, H.; TAOKA, H.; TULLBERG, H.; UUSITALO, M. A.; TIMUS, B.; FALLGREN, M. Scenarios for 5G mobile and wireless communications: the vision of the METIS project. *IEEE Communications Magazine*, v. 52, n. 5, p. 26–35, May 2014. Cited on page 16.

2 CISCO. *Cisco Annual Internet Report (2018-2023) White Paper*. 2020. 35 p. Available: ⟨https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html⟩. Cited on page 16.

3 LEIGHTON, T. Can the internet keep up with the surge in demand? *The Akamai Blog*, Apr. 2020. Accessed: Feb. 16, 2021. Available: ⟨https://blogs.akamai.com/2020/04/can-the-internet-keep-up-with-the-surge-in-demand.html⟩. Cited on page 16.

4 GRANRYD, M. How governments and mobile operators are easing network congestion during the COVID-19 crisis. *World Economic Forum*, Apr. 2020. Accessed: Feb. 16, 2021. Available: ⟨https://www.weforum.org/agenda/2020/04/how-governments-and-mobile-operators-are-easing-network-congestion-during-the-covid-19-crisis/⟩. Cited on page 16.

5 O'SHEA, T.; HOYDIS, J. An introduction to deep learning for the physical layer. *IEEE Transactions on Cognitive Communications and Networking*, v. 3, n. 4, p. 563–575, Oct. 2017. Cited 4 times on page(s) 17, 24, 25, and 26.

6 POPOVSKI, P.; BRAUN, V.; MAYER, H.-P.; FERTL, P.; REN, Z.; GONZALES-SERRANO, D.; STRÖM, E.; SVENSSON, T.; TAOKA, H.; AGYAPONG, P.; BENJEBBOUR, A.; ZIMMERMANN, G.; MEINILÄ, J.; YLITALO, J.; JÄMSÄ, T.; KYÖSTI, P.; DIMOU, K.; FALLGREN, M.; SELÈN, Y.; TIMUS, B.; TULLBERG, H.; SCHELLMANN, M.; WU, Y.; SCHUBERT, M.; KANG, D. H.; MARKENDAHL, J. I.; BECKMAN, C.; UUSITALO, M.; YILMAZ, O.; WIJTING, C.; LI, Z.; MARSCH, P.; PAWLAK, K.; VIHRIALA, J.; GOURAUD, A.; JEUX, S.; BOLDI, M.; DELL'AERA, G. M.; MELIS, B.; SCHOTTEN, H.; SPAPIS, P.; KALOXYLOS, A.; CHATZIKOKOLAKIS, K. *Scenarios, requirements and KPIs for 5G mobile and wireless system*. 2013. 84 p. Available: ⟨https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.1_v1.pdf⟩. Cited on page 17.

7 BANA, A.; CARVALHO, E. de; SORET, B.; ABRÃO, T.; MARINELLO, J. C.; LARSSON, E. G.; POPOVSKI, P. Massive MIMO for internet of things (IoT) connectivity. *Physical Communication*, v. 37, p. 100859, Sept. 2019. Cited on page 18.

8 CHEN, X.; NG, D. W. K.; YU, W.; LARSSON, E. G.; AL-DHAHIR, N.; SCHOBER, R. Massive access for 5G and beyond. *IEEE Journal on Selected Areas in Communications*, v. 39, n. 3, p. 615–637, Mar. 2021. Cited 2 times on page(s) 19 and 22.

9 LARSSON, E. G.; EDFORS, O.; TUFVESSON, F.; MARZETTA, T. L. Massive MIMO for next generation wireless systems. *IEEE Communications Magazine*, v. 52, n. 2, p. 186–195, Feb. 2014. Cited on page 19.

10 MARTÍNEZ, À. O.; CARVALHO, E. D.; Nielsen, J. Ø. Towards very large aperture massive MIMO: A measurement based study. In: *2014 IEEE Globecom Workshops (GC Wkshps)*. 2014. p. 281–286. Cited on page 19.

11 ZHOU, Z.; GAO, X.; FANG, J.; CHEN, Z. Spherical wave channel and analysis for large linear array in LoS conditions. In: *2015 IEEE Globecom Workshops (GC Wkshps)*. 2015. p. 1–6. Cited on page 19.

12 CARVALHO, E. de; ALI, A.; AMIRI, A.; ANGJELICHINOSKI, M.; HEATH, R. W. Non-stationarities in extra-large-scale massive MIMO. *IEEE Wireless Communications*, v. 27, n. 4, p. 74–80, Aug. 2020. Cited on page 20.

13 LI, X.; ZHOU, S.; BJÖRNSON, E.; WANG, J. Capacity analysis for spatially non-wide sense stationary uplink massive MIMO systems. *IEEE Transactions on Wireless Communications*, v. 14, n. 12, p. 7044–7056, Aug. 2015. Cited on page 20.

14 ALI, A.; CARVALHO, E. D.; HEATH, R. W. Linear receivers in non-stationary massive MIMO channels with visibility regions. *IEEE Wireless Communications Letters*, v. 8, n. 3, p. 885–888, Feb. 2019. Cited on page 20.

15 LI, K.; SHARAN, R. R.; CHEN, Y.; GOLDSTEIN, T.; CAVALLARO, J. R.; STUDER, C. Decentralized baseband processing for massive MU-MIMO systems. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, v. 7, n. 4, p. 491–507, Nov. 2017. Cited on page 20.

16 SÁNCHEZ, J. R.; ALEGRÍA, J. V.; RUSEK, F. Decentralized massive MIMO systems: Is there anything to be discussed? In: *2019 IEEE International Symposium on Information Theory*. 2019. p. 787–791. Cited on page 20.

17 MUELLER, A.; KAMMOUN, A.; BJÖRNSON, E.; DEBBAH, M. Linear precoding based on polynomial expansion: reducing complexity in massive MIMO. *EURASIP Journal on Wireless Communications and Networking*, v. 1, n. 63, p. 1687–1499, Feb. 2016. Cited on page 20.

18 FODOR, G.; RAJATHEVA, N.; ZIRWAS, W.; THIELE, L.; KURRAS, M.; GUO, K.; TOLLI, A.; SØRENSEN, J. H.; CARVALHO, E. de. An overview of massive MIMO technology components in METIS. *IEEE Communications Magazine*, v. 55, n. 6, p. 155–161, June 2017. Cited on page 20.

19 GARCIA-RODRIGUEZ, A.; MASOUROS, C.; RULIKOWSKI, P. Reduced switching connectivity for large scale antenna selection. *IEEE Transactions on Communications*, v. 65, n. 5, p. 2250–2263, Feb. 2017. Cited 2 times on page(s) 20 and 31.

20 AMIRI, A.; MANCHÓN, C. N.; CARVALHO, E. de. Deep learning based spatial user mapping on extra large MIMO arrays. *arXiv. 2002.00474*, February 2020. Cited on page 20.

21 MARINELLO, J. C.; ABRÃO, T.; AMIRI, A.; CARVALHO, E. de; POPOVSKI, P. Antenna selection for improving energy efficiency in XL-MIMO systems. *IEEE Transactions on Vehicular Technology*, v. 69, n. 11, p. 13305–13318, Sept. 2020. Cited on page 20.

22 AMIRI, A.; REZAIE, S.; MANCHON, C. N.; CARVALHO, E. de. Distributed receivers for extra-large scale MIMO arrays: A message passing approach. *arXiv. 2007.06930*, July 2020. Cited on page 20.

23 EPHREMIDES, A.; HAJEK, B. Information theory and communication networks: an unconsummated union. *IEEE Transactions on Information Theory*, v. 44, n. 6, p. 2416–2434, 1998. Cited on page 21.

24 BERTSEKAS, D.; GALLAGER, R. *Data Networks*. 2. ed. Englewood Cliffs, NJ, USA: Prentice Hall, 1992. Cited on page 21.

25 FRANKISH, K.; RAMSEY, W. M. *The Cambridge Handbook o Artificial Intelligence*. 1. ed. Cambridge, UK: Cambridge University Press, 2014. Cited on page 22.

26 RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: A Modern Approach*. 4. ed. Hoboken, NJ, USA: Pearson, 2021. Cited on page 22.

27 BJÖRNSON, E.; SANGUINETTI, L.; WYMEERSCH, H.; HOYDIS, J.; MARZETTA, T. L. Massive MIMO is a reality—what is next?: Five promising research directions for antenna arrays. *Digital Signal Processing*, v. 94, p. 3–20, Nov. 2019. Cited 2 times on page(s) 22 and 25.

28 EIBEN, A. E.; SMITH, J. E. *Introduction to Evolutionary Computing*. 2. ed. Berlin, DE: Springer, 2015. Cited on page 23.

29 MONTERO, E.; RIFF, M.-C.; NEVEU, B. A beginner's guide to tuning methods. *Applied Soft Computing*, v. 17, p. 39–51, Apr. 2014. Cited 2 times on page(s) 23 and 24.

30 GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. 1. ed. Cambridge, MA, USA: MIT Press, 2016. Cited 3 times on page(s) 24, 25, and 26.

31 MURPHY, K. P. *Machine Learning: A Probabilistic Perspective*. 1. ed. Cambridge, MA, USA: MIT Press, 2012. Cited on page 25.

32 SUTTON, R. S.; BARTO, A. G. *Reinforcement learning: an introduction*. 2. ed. Cambridge, MA, USA: MIT Press, 2018. Cited on page 25.

33 KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv. 1412.6980*, Jan. 2017. Cited on page 26.

34 GAO, X.; EDFORS, O.; TUFVESSON, F.; LARSSON, E. G. Massive MIMO in real propagation environments: Do all antennas contribute equally? *IEEE Transactions on Communications*, v. 63, n. 11, p. 3917–3928, July 2015. Cited on page 31.

35 TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, v. 58, n. 1, p. 267–288, Jan. 1996. Cited on page 35.

36 DONOHO, D. L.; MALEKI, A.; MONTANARI, A. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, v. 106, n. 45, p. 18914–18919, Nov. 2009. Cited 2 times on page(s) 35 and 41.

37 CHU, D. Polyphase codes with good periodic correlation properties (corresp.). *IEEE Transactions on Information Theory*, v. 18, n. 4, p. 531–532, July 1972. Cited on page 35.

38 QIN, Z.; SCHEINBERG, K.; GOLDFARB, D. Efficient block-coordinate descent algorithms for the group lasso. *Mathematical Programming Computation*, v. 5, n. 2, p. 143–169, Mar. 2013. Cited on page 41.

39 FENGLER, A.; HAGHIGHATSHOAR, S.; JUNG, P.; CAIRE, G. Non-bayesian activity detection, large-scale fading coefficient estimation, and unsourced random access with a massive MIMO receiver. *IEEE Transactions on Information Theory*, v. 67, n. 5, p. 2925–2951, Mar. 2021. Cited on page 41.

40 WRIGHT, S. J. Coordinate descent algorithms. *Mathematical Programming*, v. 151, n. 1, p. 3–34, Mar. 2015. Cited on page 41.

41 LIU, L.; YU, W. Massive connectivity with massive MIMO—Part I: Device activity detection and channel estimation. *IEEE Transactions on Signal Processing*, v. 66, n. 11, p. 2933–2946, Mar. 2018. Cited on page 41.

42 ZHANG, W.; LI, S.; CUI, Y. Jointly sparse support recovery via deep auto-encoder with applications in MIMO-based grant-free random access for mMTC. In: *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. 2020. p. 1–5. Cited on page 45.

43 CUI, Y.; LI, S.; ZHANG, W. Jointly sparse signal recovery and support recovery via deep learning with applications in MIMO-based grant-free random access. *IEEE Journal on Selected Areas in Communications*, v. 39, n. 3, p. 788–803, Aug. 2020. Cited on page 46.

44 LIU, J.; AGIWAL, M.; QU, M.; JIN, H. Online control of preamble groups with priority in massive IoT networks. *IEEE Journal on Selected Areas in Communications*, v. 39, n. 3, p. 700–713, Mar. 2021. Cited on page 46.

APPENDIX A – Quasi-Distributed Antenna Selection for Spectral Efficiency Maximization in Subarray Switching XL-MIMO Systems

Quasi-Distributed Antenna Selection for Spectral Efficiency Maximization in Subarray Switching XL-MIMO Systems

João Henrique Inacio de Souza, *Graduate Student Member, IEEE*, Abolfazl Amiri, Taufik Abrão, *Senior Member, IEEE*, Elisabeth de Carvalho, *Senior Member, IEEE*, and Petar Popovski, *Fellow, IEEE*

***Abstract*—In this paper, we consider the downlink (DL) of a zero-forcing (ZF) precoded extra-large scale massive MIMO (XL-MIMO) system. The base-station (BS) operates with limited number of radio-frequency (RF) transceivers due to high cost, power consumption and interconnection bandwidth associated to the fully digital implementation. The BS, which is implemented with a subarray switching architecture, selects groups of active antennas inside each subarray to transmit the DL signal. This work proposes efficient resource allocation (RA) procedures to perform joint antenna selection (AS) and power allocation (PA) to maximize the DL spectral efficiency (SE) of an XL-MIMO system operating under different loading settings. Two metaheuristic RA procedures based on the genetic algorithm (GA) are assessed and compared in terms of performance, coordination data size and computational complexity. One algorithm is based on a quasi-distributed methodology while the other is based on the conventional centralized processing. Numerical results demonstrate that the quasi-distributed GA-based procedure results in a suitable trade-off between performance, complexity and exchanged coordination data. At the same time, it outperforms the centralized procedures with appropriate system operation settings.**

***Index Terms*—Extra-large scale massive MIMO (XL-MIMO), antenna selection (AS), resource allocation (RA), genetic algorithm (GA), distributed signal processing.**

I. Introduction

THE benefits of adopting a high number of antennas at the base-station (BS) have attracted the interest on the massive MIMO transceiver design for the multi-antenna wireless communications systems beyond the fifth generation (B5G) and of the sixth generation (6 G). The main advantages are the large array gain, inter-channel orthogonality and channel hardening. Also, increasing the number of antenna elements can enhance the cell coverage, improving the quality-of-service (QoS) of the border-cell users [1].

When the BS array attains extreme physical dimensions to support crowded scenario locations, such as airports and large shopping malls, the system is classified as extra-large scale massive MIMO (XL-MIMO) [2]. The XL-MIMO array provides the benefits of massive MIMO with additional beam-forming resolution due to the large array aperture [3]. The XL-MIMO array is characterized by key changes in the electromagnetic propagation conditions when compared to the conventional spatial stationary massive MIMO regime. The first property is the spherical wavefront propagation feature for the received signal due to the distance between the BS and the users being less than the Rayleigh distance [4]. Second, each cluster of scatterers sees only a portion of the array. Thus, the transmitted signal by each user reaches a small group of antennas, which comprises the visibility region (VR) of this user [2]. Additionally, the different propagation paths experienced along the array result in variations on the average received power. Results in [5], [6] demonstrate that the spatial non-stationarities produced by these two properties limit the performance of the system in terms of spectral efficiency (SE) unless an appropriated signal processing technique is applied.

Despite the benefits of high numbers of antennas, the XL-MIMO scenario imposes challenges for transceiver design. The first of them is the high cost and power consumption of fully digital implementations, which require one radio-frequency (RF) transceiver per antenna element [7], [8]. In addition, adopting a large number of antennas demands a high interconnection bandwidth to transmit the baseband data throughout the links to the BS processing unit. This turns into a serious implementation bottleneck, since the required bandwidth can not be handled by the current radio interfaces [9], [10]. Lastly, handling the complexity of signal processing techniques is a relevant issue, since the number of executed operations in linear detectors, such as zero-forcing (ZF) and minimum mean-squared error (MMSE), scales with the number of antennas [11].

In order to design practical BS architectures, one can limit the number of RF transceivers to cope with the cost constraints. The

Manuscript received February 17, 2021; revised May 13, 2021; accepted May 13, 2021. Date of publication May 18, 2021; date of current version July 20, 2021. This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001, in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 310681/2019-7, and in part by the Danish Council for Independent Research under Grant DFF-701700271. The review of this article was coordinated by Dr. Maged Elkashlan. *(Corresponding author: Taufik Abrao).*

João Henrique Inacio de Souza and Taufik Abrão are with the Electrical Engineering Department, State University of Londrina, PR 86057-970, Brazil (e-mail: joaohis@outlook.com; taufik@uel.br).

Abolfazl Amiri, Elisabeth de Carvalho, and Petar Popovski are with the Department of Electronic Systems, Technical Faculty of IT and Design, Aalborg University, DK-9220 Aalborg, Denmark (e-mail: aba@es.aau.dk; edc@es.aau.dk; petarp@es.aau.dk).

Digital Object Identifier 10.1109/TVT.2021.3081462

implementation with a limited the number of RF transceivers can benefit from the large array by adopting techniques such as antenna selection (AS) and hybrid precoding. Often, hybrid precoding design is associated with the solution of intricate optimization problems [12]. In addition, the commonly employed analog phase shifters are more expensive and consume more power than conventional on-off switches [8]. For these reasons, combining the AS procedures with linear precoding designs result in attainable strategies aiming at robust and effective implementations. Different approaches and tools can be adopted to perform AS, such as convex optimization [7], [13], [14], greedy heuristics [7], [15], machine learning [16] and metaheuristics [17]–[20].

One strategy to combat the problem of high interconnection bandwidth is to use hierarchical architectures. Adding multiple processing units to handle small groups of antennas and choosing the right signal processing methods can reduce significantly the amount of exchanged information in the regime of asymptotic number of antennas, as discussed in [9], [10]. However, the coordination of such processing units to perform different signal processing and resource allocation (RA) tasks constitutes a big challenge. In addition, many of these activities rely on the knowledge of fully reliable channel state information (CSI), which is hard to attain due to the high array dimensions. Many works on channel estimation [21], precoding and data detection [9], [10], [22]–[25] in massive and XL-MIMO consider distributed pre-processing at local nodes. However, studies on the distributed RA strategies, mainly involving AS, are scarce.

The signal processing complexity is an important concern in XL-MIMO due to the high number of antenna elements. However, differently from the conventional massive MIMO, the XL-MIMO can benefit from the spatial non-stationarities adopting local signal processing strategies to treat the signals inside the VRs at the BS subarrays with reduced complexity [22], [24].

A. Literature Review

AS strategies for MIMO systems are extensively discussed in the literature. One AS algorithm to improve capacity in low rank matrix channels on point-to-point MIMO was first introduced in [26]. Later, the capacity distribution of systems with receive AS has been derived in [27]. These results were extended to massive MIMO regime in [28] and [29]. In these papers, the authors derived capacity bounds for systems with transmit and receive AS, respectively.

The authors in [13], [14] proposed AS procedures respectively for the channel capacity and downlink (DL) sum-capacity maximization based on the convex optimization framework. One technique based on the branch-and-bound algorithm is used in [8]. Considering linearly-precoded systems, the problems of AS for SE and sum-SINR maximization are addressed respectively in [15], [30]. Differently, the work in [31] analyzed one joint AS and power allocation (PA) procedure in a system with spatially distributed antennas. The proposed procedure runs at each antenna with side-information shared within its neighborhood. Besides, AS considering limited connections in the RF transceivers switching matrices is examined in [7].

On the other hand, there are only a few works that consider the AS problem for the XL-MIMO systems. A spatial users mapping procedure to maximize SE implemented with convolutional neural networks (CNN) is proposed in [16]. The aim is to determine each effective subarray window to precode the users signals using ZF. Results demonstrate that the CNN-based procedure achieves SE values comparable to the optimal mapping algorithm. In [17], several transmit AS procedures to maximize the energy efficiency (EE) from the long-term fading coefficients are proposed. Asymptotic SINR expressions for the received signal with AS are derived. Since the derived optimization problem is NP-hard, three of the proposed procedures are implemented by metaheuristic techniques, one being the genetic algorithm (GA). The GA is a powerful evolutionary metaheuristic that was used in different contexts to solve AS problems, as it is considered in [18]–[20].

B. Contribution

Motivated by the benefits of large numbers of antennas at the BS and the restricted number of RF transceivers, this work examines the joint AS and PA problem on the DL of a linearly-precoded XL-MIMO system. Differently from other papers adopting AS strategy, a distributed BS signal processing architecture is considered and the AS procedures are characterized in terms of the exchanged information between the processing nodes. Furthermore, we extend part of the results of [17] with the proposition of AS algorithms for XL-MIMO that use the short-term fading coefficients instead of the long-term ones. Additionally, we address the problem of joint AS and PA in XL-MIMO subarrays using a decentralized RA algorithm. The proposed RA algorithm uses the Sherman-Morrison-Woodbury (SMW) formula to perform optimal power allocation (OPA) and AS in a decentralized fashion.

The BS is constituted by multiple non-overlapping subarrays with dedicated remote processing units (RPUs), which perform independently channel estimation, precoding calculation and RA, mainly AS and PA. Each subarray is equipped with a fixed number of antenna elements and RF transceivers. Using the ZF precoding, the optimization goal is to maximize the SE subjected to the constraints of subarrays connections and maximum transmitted power.

The contribution of this work is fourfold. *i*) Description of a distributed transceiver design for XL-MIMO based on a subarray switching architecture; *ii*) proposition of a centralized procedure based on the evolutionary heuristic GA to perform joint AS and PA to maximize the SE with subarray connection and maximum transmitted power constraints; *iii*) proposition of a distributed version of the GA procedure for joint AS and PA which achieves performance tight to the centralized one but with low-size coordination data and less number of executed operations; *iv*) extensive analysis of the proposed procedures in terms of number of symbols for training, coordination data size and number of floating point operations per second (flops).

The numerical results corroborate the GA-based procedures in achieving high performance, specifically in crowded XL-MIMO applications. Additionally, the decentralized GA version offers a good trade-off between performance, number of operations and

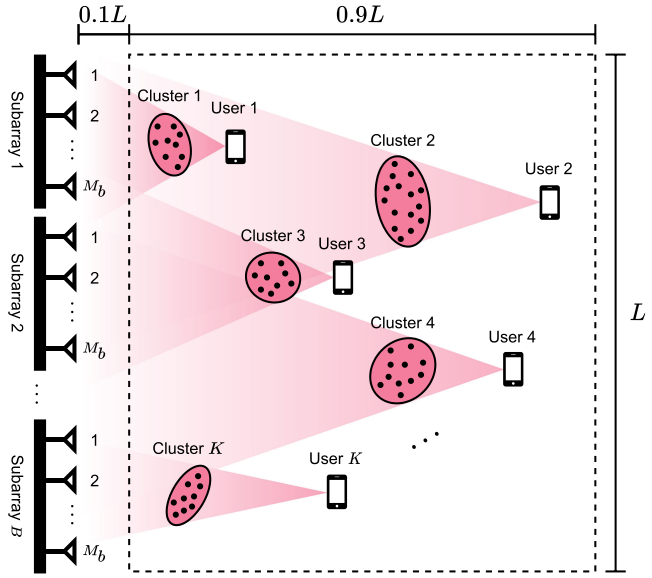


Fig. 1. XL-MIMO system deployed inside a square cell with size L. The BS is a ULA with M antennas divided into B subarrays of M_b antennas each one. The K users are randomly distributed at a distance in the range $(0.1^{-}L, L)$ from the array.

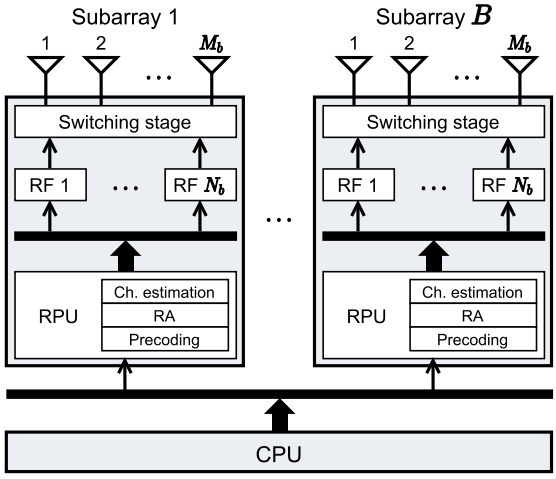


Fig. 2. Diagram of the BS architecture for DL. The BS array is composed by B subarrays containing M_b antennas, N_b RF transceivers and one RPU. Additionally, the BS has a CPU for subarrays coordination.

coordination data size, outperforming the centralized procedures by adopting proper settings.

The rest of the paper is organized as follows. In Section II is described the system model, including the distributed subarrays processing at the BS. Next, in Section III are described the centralized and distributed GA-based optimization procedures for joint AS and PA in XL-MIMO systems, while Section IV discusses two feasible AS procedures adopted as a result of decoupling the joint AS and PA optimization problem. Section V examines the complexity of the proposed algorithms. Extensive numerical results are discussed in Section VI. Final comments and conclusions are provided in Section VII.

C. Notation

Boldface small $\mathbf{a}$ and capital $\mathbf{A}$ letters represent respectively vectors and matrices. Capital calligraphic letters $\mathcal{A}$ represent finite sets, and $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. $\mathbf{I}_n$ denotes the identity matrix of size n. $\{\cdot\}^T$ and $\{\cdot\}^H$ denote respectively the transpose and the conjugate transpose operators. $\text{diag}(\cdot)$, $\text{tr}(\cdot)$ and $\det(\cdot)$ denote respectively the diagonal matrix, trace and determinant operators. $-plx - left - lceil\cdot\rceil$ denotes the greatest integer operator. $\binom{n}{k}$ denotes the binomial coefficient. $\mathcal{CN}(\mu, \sigma^2)$ is a circularly symmetric complex Gaussian distribution with mean μ and variance σ^2. $\mathbb{E}[\cdot]$ denotes the expectation operator.

II. System Model

Consider the DL of a narrow-band multi-user XL-MIMO system with the BS equipped with M antennas and N RF transceivers serving K single-antenna users, as is depicted in Fig. 1. During the DL, the BS uses η_{tr} symbols to perform channel estimation and η_{data} symbols to transmit the payload. We assume that the time interval used to send the total DL symbols $\eta_{\text{DL}} = \eta_{\text{tr}} + \eta_{\text{data}}$ is less than the channel coherence time.

The array in the BS is composed of B independent subarrays, each with M_b antennas and $N_b < M_b$ RF transceivers. The subarrays are equipped with a RPU to perform, in a distributed way, channel estimation, precoding calculation and RA tasks, specially AS and PA procedures. In addition, the BS has a central processing unit (CPU) to coordinate the subarrays operation. Fig. 2 depicts all the described BS blocks.

Assumption 1 (Subarray switching stage): A flexible switching stage is implemented in each XL subarray. This stage allows every antenna of the subarray i to connect to any RF transceiver of it. Results in [7] demonstrate that partially connected architectures introduce lower insertion loss than fully-flexible matrices, which allows the connection of any antenna in the entire array to any RF transceiver.

We assume that each subarray has perfect knowledge of the channel coefficients associated to its antennas. See [21] for details on channel acquisition in distributed signal processing architectures. Besides, we deploy the ZF precoder to decode signals in each subarray. We adopt the technique in [21] to calculate the ZF precoder with low interconnection traffic, splitting the computations between the RPUs and the CPU.

A. Channel Model

In the XL-MIMO scenario, spatial non-stationarities arise due to the large array physical dimensions and number of antenna elements. Such non-stationarities are addressed in the adopted channel model as the variation of the mean received power along the array, as in [17], [22]. The path-loss coefficient associated to the BS antenna m and the user k is defined as

$$\beta_{m,k} = q_0 d_{m,k}^{-\kappa} \tag{1}$$

where q_0 is the path-loss attenuation at a reference distance, $d_{m,k}$ is the distance between the antenna m and the user k and κ is the path-loss exponent.

Let $\mathbf{R}_k \in \mathbb{C}^{M\times M}$, $\mathbf{R}_k = \text{diag}([\beta_{1,k} \ \cdots \ \beta_{M,k}]^T)$ be the matrix with the long-term fading coefficients of the user k. The channel vector of the user k is defined as

$$\mathbf{h}_k = \mathbf{R}_k^{\frac{1}{2}}\mathbf{h}'_k \tag{2}$$

where $\mathbf{h}'_k \in \mathbb{C}^{M\times 1}$, $\mathbf{h}'_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is the short-term fading vector. From the users channel vectors, the channel matrix $\mathbf{H} \in \mathbb{C}^{M\times K}$ is defined as

$$\mathbf{H} = \begin{bmatrix}\mathbf{h}_1 & \cdots & \mathbf{h}_K\end{bmatrix} = \begin{bmatrix}\underline{\mathbf{h}}_1^T & \cdots & \underline{\mathbf{h}}_M^T\end{bmatrix}^T \tag{3}$$

considering $\underline{\mathbf{h}}_m \in \mathbb{C}^{1\times K}$ as the channel vector with the coefficients associated to the antenna m.

During the DL, the BS activates a group of antennas represented by the set $\mathcal{S} \subseteq \{1,\ldots,M\}$ such that $|\mathcal{S}| \leq N$. A partition of the set $\mathcal{S}$, *i.e.* $\{\mathcal{S}_b\}, \forall b = 1,\ldots,B$, contains the index of the selected antennas in the subarray b. This set is defined such that $|\mathcal{S}_b| \leq N_b \ \forall b$, meeting the adopted subarray structure. The equivalent channel matrix of the active antennas is defined as a row-wise submatrix of $\mathbf{H}$, $\mathbf{H}_{\mathcal{S}} \in \mathbb{C}^{|\mathcal{S}|\times K}$. Similarly, the matrix $\mathbf{H}_{\mathcal{S}_b} \in \mathbb{C}^{|\mathcal{S}_b|\times K}$ contains only the channel vectors related to the active antennas in the subarray b.

Let $D_m \in \{0,1\}, \ \forall m = 1,\ldots,M$ be an indicator equal to 1 if the antenna m is active during the DL and 0 otherwise. These indicators form the diagonal matrix $\mathbf{D} = \text{diag}([D_1 \ \cdots \ D_M]^T)$. During the precoding and SE computations, it is required to calculate the matrix product $\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}}$ of the active antennas channel matrix. Intended to enable this computation by the distributed signal processing architecture, the Gramian matrix is defined as in the following.

Remark 1 (Gramian matrix): Let $\mathbf{G}_m = \underline{\mathbf{h}}_m^H\underline{\mathbf{h}}_m, \ \forall m = 1,\ldots,M$ be the Gramian matrix associated with the BS antenna m. The set $\mathcal{M}_b$ is defined for $b = 1,\ldots,B$ as the group of antennas in the subarray b. The Gramian matrix associated to the b-th subarray includes only the active antennas inside it, and it can be written as

$$\mathbf{G}_{\mathcal{S}_b} = \mathbf{H}_{\mathcal{S}_b}^H\mathbf{H}_{\mathcal{S}_b} = \sum_{m\in\mathcal{M}_b} D_m\mathbf{G}_m \tag{4}$$

Similarly, the array Gramian matrix considering only the active antennas is defined as

$$\mathbf{G}_{\mathcal{S}} = \mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}} = \sum_{m=1}^{M} D_m\mathbf{G}_m \tag{5}$$

An upper bound for the system performance considering the active antennas in the set $\mathcal{S}$, namely the DL sum-capacity, is calculated by [14]:

$$\begin{aligned} C_{\text{DPC}} &= \max_{\mathbf{P}} \log_2 \det\left(\mathbf{I}_K + \frac{1}{\sigma_z^2}\mathbf{P}\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}}\right) \\ &= \max_{\mathbf{P}} \log_2 \det\left(\mathbf{I}_K + \frac{1}{\sigma_z^2}\mathbf{P}\mathbf{G}_{\mathcal{S}}\right) \end{aligned} \tag{6}$$

where σ_z^2 is the additive noise power, while $\mathbf{P} = \text{diag}([p_1 \ \cdots \ p_K])$ denotes the matrix with the allocated power for each user. The powers $p_k, \ \forall k = 1,\ldots,K$ are defined in order to meet the total power constraint $\sum_{k=1}^K p_k = P_{\max}$. The DL sum-capacity is achieved by the *dirty paper coding* (DPC) precoder, which has prohibitive high-complexity for practical implementations.

B. Downlink Signal

The data signal transmitted by the BS is defined as $\mathbf{x} \in \mathbb{C}^{|\mathcal{S}|\times 1}$,

$$\mathbf{x} = \mathbf{F}\mathbf{P}^{\frac{1}{2}}\mathbf{s} \tag{7}$$

where $\mathbf{F} \in \mathbb{C}^{|\mathcal{S}|\times K}$ denotes the ZF precoding matrix, calculated by

$$\begin{aligned} \mathbf{F} &= \mathbf{H}_{\mathcal{S}}\left(\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}}\right)^{-1} \\ &= \mathbf{H}_{\mathcal{S}}\mathbf{G}_{\mathcal{S}}^{-1} \end{aligned} \tag{8}$$

$\mathbf{s} = [s_1 \ \cdots \ s_K]^T$ denotes the vector of modulated data symbols such that $\mathbb{E}[\|s_k\|_2^2] = 1, \ \forall k = 1,\ldots,K$ and $\mathbb{E}[s_k^* s_{k'}] = 0, \ \forall k \neq k'$. The allocated powers in (7) are calculated in order to meet the following power constraint

$$\text{tr}\left[\mathbf{P}\left(\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}}\right)^{-1}\right] = \text{tr}\left(\mathbf{P}\mathbf{G}_{\mathcal{S}}^{-1}\right) = P_{\max} \tag{9}$$

Therefore, the entries of $\mathbf{P}$ depend on the active antennas set $\mathcal{S}$ and the PA policy.

The signal received by the users in the DL is defined as $\mathbf{y} \in \mathbb{C}^{K\times 1}$,

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_{\mathcal{S}}^H\mathbf{F}\mathbf{P}^{\frac{1}{2}}\mathbf{s} + \mathbf{z} \\ &= \mathbf{P}^{\frac{1}{2}}\mathbf{s} + \mathbf{z} \end{aligned} \tag{10}$$

where $\mathbf{z} \in \mathbb{C}^{K\times 1}$, $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_z^2\mathbf{I}_K)$ denotes the additive noise vector.

Given the ZF precoding design, the system SE is calculated by

$$\text{SE} = \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k}{\sigma_z^2}\right) \tag{11}$$

which is equivalent to the SE of K independent Gaussian channels with received signal-to-noise ratio (SNR) equal to $p_k/\sigma_z^2 \ \forall k$.

C. Optimal Power Allocation (OPA) Policy

The OPA policy is the one that solves the problem of maximizing the system SE at (11), subjected to the maximum power constraint in (9):

$$\underset{\mathbf{P}}{\text{maximize}} \quad \text{SE} = \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k}{\sigma_z^2}\right) \tag{12a}$$

$$\text{subject to} \quad \text{tr}\left[\mathbf{P}(\mathbf{H}_{\mathcal{S}}^H\mathbf{H}_{\mathcal{S}})^{-1}\right] \leq P_{\max} \tag{12b}$$

$$p_k \geq 0, \ \forall k = 1,\ldots,K \tag{12c}$$

The optimization problem in (12) is equivalent to the well-known PA problem on independent Gaussian channels. It has an analytical closed-form solution derived by the Lagrange multipliers method (water filling solution). The optimal power

distribution is calculated by [32]:

$$p_k = \left(\mu \left[(\mathbf{H}_{\mathcal{S}}^H \mathbf{H}_{\mathcal{S}})^{-1}\right]_{k,k}^{-1} - \sigma_z^2\right)^+ \tag{13}$$

where $(x)^+ = \max(x, 0)$ and μ is a constant calculated by

$$\mu = \frac{1}{K}\left\{P_{\max} + \sigma_z^2 \text{tr}\left[(\mathbf{H}_{\mathcal{S}}^H \mathbf{H}_{\mathcal{S}})^{-1}\right]\right\} \tag{14}$$

If $p_k = 0$ for some user k, the PA problem including this user is not feasible. For this reason, the k-th user is deactivated and the power distribution is recalculated considering only the group of the remaining active users. This process must be repeated until a group of users which results in a feasible solution is found.

III. Algorithm for Joint Antenna Selection and Power Allocation

The problem of jointly selecting the antenna-elements of the BS and allocating appropriate power amounts to maximizing the ZF SE given the constraints of maximum RF transceivers, subarray connections, and maximum power is formulated as

$$\underset{\mathbf{D},\mathbf{P}}{\text{maximize}} \;\; \text{SE} = \sum_{k=1}^{K} \log_2\left(1 + \frac{p_k}{\sigma_z^2}\right) \tag{15a}$$

$$\text{subject to} \sum_{m \in \mathcal{M}_b} D_m \leq N_b, \quad b = 1, \ldots, B \tag{15b}$$

$$\text{tr}\left[\mathbf{P}(\mathbf{H}^H \mathbf{D} \mathbf{H})^{-1}\right] \leq P_{\max} \tag{15c}$$

$$D_m \in \{0, 1\}, \qquad m = 1, \ldots, M \tag{15d}$$

$$p_k \geq 0, \qquad k = 1, \ldots, K \tag{15e}$$

The objective function in (15a) is the system SE. The constraints (15b) are the subarray connections constraints, which allow the activation of a maximum of N_b RF transceivers in each subarray. Also, the constraint (15c) ensures that the maximum transmitted power is equal to or less than $P_{\max}$. Moreover, the constraints (15d) and (15e) define respectively the binary antenna association variables and non-negative allocated powers.

Since $\mathbf{D}$ is binary constrained, the problem (15) constitutes a non-convex combinatorial optimization problem. One approach to solve (15) comprises two steps: firstly, determining the optimal active antennas set via exhaustive search assuming equal PA; after that, given the result $\mathbf{D}^\star$ from the exhaustive search, the allocated power matrix $\mathbf{P}^\star$ is calculated adopting the OPA policy in (13).

The AS via exhaustive search considering the activation of all the RF transceivers requires testing $\binom{M_b}{N_b}^B$ candidate solutions, a number that attains prohibitive dimensions in the XL-MIMO regime. For instance, in a system with $B = 8$ subarrays equipped with $M_b = 64$ antennas and $N_b = 32$ RF transceivers, there is a number of feasible solutions on the order of magnitude equal to 10^{146}. Testing all these solution candidates in a timely manner is impracticable. An efficient alternative to the exhaustive search is to perform a guided search along the feasible set using an intelligent metaheuristic procedure. In this way, a good quality solution can be obtained in feasible time testing only a few candidates.

TABLE I
Glossary of the Genetic Algorithm Terms

Parameter	Description
Individual	Candidate solution for the optimization problem
Population	Set of candidate solutions for the optimization problem
Offspring	Set of candidate solutions generated during an iteration
Gene	One optimization variable of the candidate solution
Chromosome	Set of optimization variables of the candidate solution
Generation	Genetic algorithm iteration
Fitness	Objective function of the optimization problem
Score	Value of the objective function for a candidate solution

A. Genetic Algorithm

One metaheuristic procedure adopted to solve many different combinatorial problems in wireless communications is the GA. This technique implements different search phases to efficiently explore the feasible set and exploit the good candidates properties in order to find promising regions in the feasible subspaces. Differently from exact optimization methods, evolutionary metaheuristics do not require convex objective functions or constraints. In addition, the execution complexity can be fitted to the available computational burden by adjusting the input parameters and number of iterations. Despite the advantages, the GA, as well as other metaheuristics, does not ensure finding the optimal solution.

In general, GA solves efficiently a variety of optimization problems [17]–[20]. Our implemented version of GA has attained near-optimum solutions for the problem (15) in a competitive time/complexity, owing to its competitive performance-complexity trade-off. Besides, the GA comprises smart mechanisms to balance exploration and exploitation of the search subspaces, as well as avoiding local optima in complex cost surfaces. Lastly, unlike single-state strategies, *e.g.* simulated annealing and tabu search, the GA has a readily parallelizable algorithmic structure, capable of providing fast convergence [33].

As the GA is a procedure inspired by principles of genetics and natural selection, it inherited several terms from biology. To simplify understanding, Table I contains a glossary of some common GA terms adopted throughout this work. In the following, the implemented GA procedures, phases and variables deployed to solve the problem (15) are briefly described.

Optimization variables encoding: The optimization variables of the problem (15) are the antennas state indicators D_m and the users allocated powers p_k. The powers p_k are determined by the OPA, eq. (13). Therefore, only the antennas indicators should be encoded as individuals. Thus, the antennas state indicators $D_m,\ \forall m = 1, \ldots, M$ are defined as genes and the vectors $\underline{\mathbf{d}}_{i,b} \in \{0, 1\}^{M_b \times 1}$ such that $[\underline{\mathbf{d}}_{i,b}]_m = D_m,\ \forall m \in \mathcal{M}_b,\ b = 1, \ldots, B$ containing the optimization variables w.r.t. each subarray represent the chromosomes, where i is the individual index. Every individual is defined by a vector $\mathbf{d}_i \in \{0, 1\}^{M \times 1}$,

$$\mathbf{d}_i = \begin{bmatrix} \underline{\mathbf{d}}_{i,1}^T & \cdots & \underline{\mathbf{d}}_{i,B}^T \end{bmatrix}^T = \begin{bmatrix} D_1 & \cdots & D_M \end{bmatrix}^T \tag{16}$$

Fitness function: The fitness function considered for the implementation is the ZF SE defined in (11), with the power distribution computed by the OPA policy.

Algorithm 1: Mutation Procedure.

Input: Crossover offspring $\mathcal{P}_c$, p_m, B, M_b, N_b
Output: Mutated offspring $\mathcal{P}_m$

$\mathcal{P}_m \leftarrow \emptyset$;
for $\mathbf{d}_i \in \mathcal{P}_c$ **do**
  for $b = 1 : B$ **do**
    if *rand uniform*$(0,1) \leq p_m$ **then**
      $k \leftarrow$ rand discrete uniform$(1, M_b)$;
      if $[\underline{\mathbf{d}}_{i,b}]_m == 0$ *and* $\sum_{j=1}^{M_b} [\underline{\mathbf{d}}_{i,b}]_j == N_b$ **then**
        Go to line 5;
      $[\underline{\mathbf{d}}_{i,b}]_m \leftarrow \text{flip}([\underline{\mathbf{d}}_{i,b}]_m)$;
  $\mathcal{P}_m \leftarrow \mathcal{P}_m \cup \mathbf{d}_i$;

The implemented GA contains the following phases: *a)* elitism, *b)* tournament selection, *c)* crossover and *d)* mutation. These phases require the definition of the parameters: population size N_p, number of individuals for elitism N_e, number of tournaments N_s, crossover probability p_c and mutation probability p_m. Each procedure is summarized in the sequel.

Elitism: The elitism aims to keep the best individuals of the current generation without change. At every generation, the N_e best individuals are chosen as the first individuals of the next generation. Elitism ensures that the SE obtained with the best AS indices of the GA iteration is always a non-decreasing value.

Tournament selection: During the tournament selection, the individuals are pairwise randomly compared according to their score values. The winners of the N_s tournaments become candidates for the crossover phase. The selection step compares the sets of AS indices produced at each GA iteration according to the SE achieved by them.

Crossover: The crossover phase aims to mix the chromosomes of the tournaments winners in order to obtain new solutions. This phase exploits the good properties of the current set of AS indices. Two tournament winners, named parent 1 and parent 2, are randomly selected to generate two new individuals. Each chromosome of child 1 has the probability p_c of being inherited from parent 1 and $1 - p_c$ from parent 2. Considering child 2, every chromosome has the probability p_c of being inherited from parent 2 and $1 - p_c$ from parent 1.

Mutation: The mutation phase aims to add random small changes at the offspring generated by crossover. This phase promotes the variability among the set of AS indices, exploring different regions of the feasible set. The chromosomes are mutated with probability p_m, when one random selected gene of the chromosome is flipped. To preserve the feasibility of the solutions, the mutation phase is implemented by the scheme of Algorithm 1. The set $\mathcal{P}_c$ denotes the offspring generated during the crossover, and $\mathcal{P}_m$ is the offspring after mutation.

Convergence: There are several mechanisms to check the GA convergence. Herein, the implemented algorithm has two different criteria: the maximum number of generations $T_{\max}$ and the no improvement of the best score during the last T_{stall} generations.

Algorithm 2: GA-RA.

Input: $N_p, N_e, N_s, p_c, p_m, T_{\max}, T_{\text{stall}}, B, M_b, N_b, \mathbf{H}$
Output: The best selected antennas set, $\mathbf{D}^\star$

$\mathcal{P}_0 \leftarrow \emptyset$;
$\mathcal{P}_0 \leftarrow \mathcal{P}_0 \cup \text{N-AS}(\mathbf{H})$ *(Section IV-B)*;
for $i = 1 : N_p - 1$ **do**
  $\mathcal{P}_0 \leftarrow \mathcal{P}_0 \cup$ rand individual();
for $t = 0 : T_{\max}$ **do**
  $\mathcal{P}_{t+1}, \mathcal{P}_s, \mathcal{P}_c \leftarrow \emptyset$;
  $\mathcal{P}_{\text{temp}} \leftarrow \mathcal{P}_t$;
  for $i = 1 : N_e$ **do** *Elitism*
    $\mathbf{d}_e \leftarrow \underset{\mathbf{d}_j}{\text{argmax}}\ \text{score}(\mathbf{d}_j),\ \mathbf{d}_j \in \mathcal{P}_{\text{temp}}$;
    $\mathcal{P}_{t+1} \leftarrow \mathcal{P}_{t+1} \cup \mathbf{d}_e$;
    $\mathcal{P}_{\text{temp}} \leftarrow \mathcal{P}_{\text{temp}} \backslash \mathbf{d}_e$;
  for $i = 1 : N_s$ **do** *Tournament selection*
    $\mathbf{d}_{s_1}, \mathbf{d}_{s_2} \leftarrow \text{rand}(\mathcal{P}_t)$;
    $\mathbf{d}_s \leftarrow \underset{\mathbf{d}_j}{\text{argmax}}\ [\text{score}(\mathbf{d}_{s_1}),\ \text{score}(\mathbf{d}_{s_2})]$;
    $\mathcal{P}_s \leftarrow \mathcal{P}_s \cup \mathbf{d}_s$;
  for $i = 1 : N_e$ **do** *Crossover*
    $\mathbf{d}_{c_1}, \mathbf{d}_{c_2} \leftarrow \text{rand}(\mathcal{P}_s)$;
    $\mathbf{d}_{o_1}, \mathbf{d}_{o_2} \leftarrow \mathbf{0}_M$
    for $j = 1 : B$ **do**
      if *rand uniform*$(0,1) \leq p_c$ **then**
        $\underline{\mathbf{d}}_{o_1,j} \leftarrow \underline{\mathbf{d}}_{c_1,j}$;
        $\underline{\mathbf{d}}_{o_2,j} \leftarrow \underline{\mathbf{d}}_{c_2,j}$;
      else
        $\underline{\mathbf{d}}_{o_1,j} \leftarrow \underline{\mathbf{d}}_{c_2,j}$;
        $\underline{\mathbf{d}}_{o_2,j} \leftarrow \underline{\mathbf{d}}_{c_1,j}$;
    $\mathcal{P}_c \leftarrow \mathcal{P}_c \cup \mathbf{d}_{o_1} \cup \mathbf{d}_{o_2}$;
  $\mathcal{P}_m \leftarrow \text{mutation}(\mathcal{P}_c)$ *(Algorithm 1)*;
  $\mathcal{P}_{t+1} \leftarrow \mathcal{P}_{t+1} \cup \mathcal{P}_m$;
  $\mathbf{d}^\star_{t+1} \leftarrow \underset{\mathbf{d}_i}{\text{argmax}}\ \text{score}(\mathbf{d}_i),\ \mathbf{d}_i \in \mathcal{P}_{t+1}$;
  if $t > T_{\text{stall}}$ **then** *Stall convergence criterion*
    $\mathbf{d}_{\text{stall}} \leftarrow \underset{\mathbf{d}_i}{\text{argmax}}\ \text{score}(\mathbf{d}_i),\ \mathbf{d}_i \in \mathcal{P}_{t-T_{\text{stall}}}$;
    if *score*$(\mathbf{d}^\star_{t+1})$ == *score*$(\mathbf{d}_{\text{stall}})$ **then**
      Break the loop;
$\mathbf{D}^\star \leftarrow \text{diag}(\mathbf{d}^\star_{t+1})$;
return $\mathbf{D}^\star$;

Algorithm 2 summarizes the implemented procedure, named *genetic algorithm for resource allocation* (GA-RA). The set $\mathcal{P}_0$ denotes the initial population, $\mathcal{P}_t$ the population of the generation t, $\mathcal{P}_s$ the winners of the tournament selection and $\mathcal{P}_{\text{temp}}$ a temporary set for the elitism phase.

B. Quasi-Distributed Genetic Algorithm

The proposed GA-RA procedure requires the entire channel matrix $\mathbf{H}$ knowledge at the CPU to compute the individuals score values. Such requirement is unfeasible in the XL-MIMO scenario due to the high bandwidth to transfer all the channel

coefficients associated to thousands of antennas to the CPU. For this reason, one solution that does not depend on the knowledge of full CSI at the CPU is preferable.

One solution to avoid the requirement of full knowledge of the $\mathbf{H}$ matrix consists of performing local AS at each subarray, considering fixed the AS indices in the other subarrays. The contribution of these fixed AS indices can be calculated previously by the CPU and transmitted to the RPUs with reduced bandwidth and processing power resources. Therefore, each subarray can selects its antennas using the GA. The proposed *quasi-distributed genetic algorithm for resource allocation* (DGA-RA) implements this concept and is presented in the following.

Analyzing the fitness function of the GA-RA procedure in (11), one can observe that it depends on the inverse of the array Gramian matrix, $\mathbf{G}_{\mathcal{S}}^{-1} = (\mathbf{H}_{\mathcal{S}}^H \mathbf{H}_{\mathcal{S}})^{-1}$. The computation of $\mathbf{G}_{\mathcal{S}}^{-1}$ can be done from the subarrays Gramian matrices by

$$\mathbf{G}_{\mathcal{S}}^{-1} = \left(\sum_{b=1}^{B} \mathbf{G}_{\mathcal{S}_b} \right)^{-1} \tag{17}$$

Therefore, the CPU can compute the inverse of the array Gramian matrix to calculate the GA-RA fitness function only with the subarrays Gramian matrices calculated locally at the RPUs. Each subarray Gramian matrix has K^2 entries, while the channel matrix has MK. Therefore, calulating the contribution of the selected antennas at the CPU using the Gramian matrix strategy requires less bandwidth than by using the centralized strategy if $BK^2 < MK$ holds.

Based on (17), the DGA-RA procedure operates as follows. Initially, each subarray selects an active antennas set based on a simple criterion, such as the *norm-based antenna selection* (N-AS) described in the subsection IV-B. Then, the subarrays compute their Gramian matrices based on the selected set and transmit them to the CPU. At the CPU, the array Gramian matrix is computed by (17) and transmitted back to the subarrays. Afterwards, every subarray performs local antenna selection by a GA implementation, considering that the other subarrays are fixed. To evaluate the fitness function in eq. (11), the subarrays compute the array Gramian inverse matrix adopting the SMW formula for matrix inversion, as follows.

Remark 2 (SMW formula): The SMW formula [34] gives the inverse of the matrix $(\mathbf{A} + \mathbf{U}\mathbf{V}^H)$ from $\mathbf{A}^{-1}$, $\mathbf{U}$ and $\mathbf{V}$ by computing:

$$(\mathbf{A} + \mathbf{U}\mathbf{V}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{I} + \mathbf{V}^H\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}^H\mathbf{A}^{-1} \tag{18}$$

Adopting this formulation, the array Gramian matrix can be calculated at the subarray b during the iteration n by letting

$$\mathbf{A}^{-1} = \left(\mathbf{G}_{\mathcal{S}}^{(n-1)}\right)^{-1}, \tag{19}$$

$$\mathbf{U} = \left[-\left(\mathbf{H}_{\mathcal{S}_b}^{(n-1)}\right)^H \quad \left(\mathbf{H}_{\mathcal{S}_b}^{(n)}\right)^H \right], \tag{20}$$

$$\mathbf{V}^H = \begin{bmatrix} \mathbf{H}_{\mathcal{S}_b}^{(n-1)} \\ \mathbf{H}_{\mathcal{S}_b}^{(n)} \end{bmatrix}, \tag{21}$$

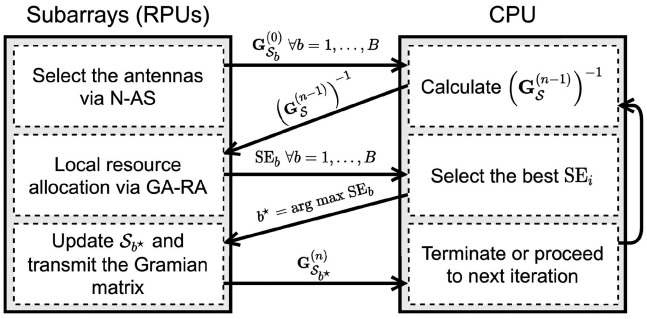


Fig. 3. Proposed DGA-RA procedure steps with coordination between the CPU and the RPUs. The superscript (n) denotes the n-th iteration.

where the superscript (n) denotes the variable during the n-th iteration of the DGA-RA procedure (proof in Appendix IX).

After performing local AS, each subarray transmits their achieved SE values to the CPU. The CPU updates the AS indices of the subarray that has achieved the maximum SE values at the iteration n. Then, the CPU requests the subarray Gramian matrix of the updated subarray, and recalculates the inverse of the array Gramian matrix, $(\mathbf{G}_{\mathcal{S}}^{(n)})^{-1}$. The process can be executed iteratively following the scheme depicted in Fig. 3.

The GA implemented in the DGA-RA procedure is similar to that one described in the Algorithm 2, except for some details at the optimization variables encoding and the crossover phase. About the individual encoding, the optimization variables at each subarray are reduced from M to M_b, since local AS is performed at each RPU. In addition, as the optimization variables consider only one subarray at each RPU, the individuals have two chromosomes: one represented by the first $M_b/2$ genes, and another composed by the remaining genes.

Due to this new chromosome definition, one further procedure after the crossover phase is required to preserve the feasibility of the solution. The chosen method is to deactivate antennas of individuals with more than N_b antennas in a random fashion until they become feasible.

IV. Antenna Selection Procedures

Two techniques to perform antenna selection are presented in the sequel, the DL *sum-capacity maximization antenna selection* (SCMAX-AS) and the N-AS method, proposed respectively in [14], [7]. The goal of solving only the antenna selection problem is to decouple the two RA problems associated to (15) aiming at obtaining tractable formulations.

A. Antenna Selection for DL Sum-Capacity Maximization

Firstly, we analyze the equal power allocation (EPA) strategy, *i.e.* $\mathbf{P} = \frac{P_{\max}}{K}\mathbf{I}_K$, intended to obtain a manageable optimization problem. The problem of selecting the set of active antennas in order to maximize the DL sum-capacity with the constraints of maximum number of RF transceivers and subarray connections is formulated as [14]:

$$\underset{\mathbf{D}}{\text{maximize}} \quad C_{\text{EPA}} = \log_2 \det\left(\mathbf{I}_K + \frac{P_{\max}}{K\sigma_z^2}\mathbf{H}^H\mathbf{D}\mathbf{H}\right) \tag{22a}$$

$$\text{subject to} \quad \sum_{m \in \mathcal{M}_b} D_m \leq N_b, \qquad b = 1, \ldots, B \tag{22b}$$

$$D_m \in \{0, 1\}, \qquad m = 1, \ldots, M \tag{22c}$$

Despite the concavity of the objective function in (22a) [13], the problem (22) is not convex due to the binary constraint in (22c). Hence, we define a convex relaxation of (22) by taking the variables D_m in the range $(0, 1)$. This new problem, which can be solved with convex optimization tools, has the constraint (22c) replaced by

$$0 \leq D_m \leq 1, \qquad m = 1, \ldots, M \tag{23}$$

Notice that the solution of the convex relaxation results in non-binary values for the active antenna indicators D_m, which is outside the original problem domain.

One method for performing the antenna selection by solving the convex relaxation is to activate the N_b antennas with the highest D_m values at each subarray. This procedure is named in this work as SCMAX-AS, and is followed by the OPA policy in eq. (13). This AS procedure gives near-optimal results, except for $N \ll M$ [14]. Therefore, in a XL-MIMO system where the number of available RF transceivers is much less than the array antennas, the achieved system SE with the SCMAX-AS algorithm will be sub-optimal.

B. Norm-Based Antenna Selection (N-AS)

The N-AS procedure focus on selecting the subset of N_b antennas with the highest channel vector norm values [7]. We adopt this method to initiate the population of the GA-based procedures due to its low computational cost. The N-AS method solves the optimization problem formulated as

$$\underset{\mathbf{D}}{\text{maximize}} \quad \Pi = \sum_{m=1}^{M} D_m \|\underline{\mathbf{h}}_m\|_2^2 \tag{24a}$$

$$\text{subject to} \quad \sum_{m \in \mathcal{M}_b} D_m \leq N_b, \qquad b = 1, \ldots, B \tag{24b}$$

$$D_m \in \{0, 1\}, \qquad m = 1, \ldots, M \tag{24c}$$

where the objective function consists of the sum of the squared norms of the channel vectors associated to the selected antennas. The problem (24) can be solved quickly by selecting the N_b antennas with the highest channel vector norms at each subarray. After selection, the PA is performed by the OPA policy in (13).

V. Complexity Analysis

The complexity of the presented procedures is evaluated in terms of the number of symbols required for channel acquisition, the size of the coordination data exchanged between the RPUs and the CPU, and the number of flops during execution.

A. Training

In the following, we analyze the procedures in terms of training symbols for CSI acquisition. The length of the mutually orthogonal pilot signals used to estimate the channel vectors at the BS depends on: *a*) the number of users; *b*) the number of available RF transceivers; *c*) the number of antennas at the BS.

The number of symbols to acquire the entire channel matrix, required in all the presented procedures except in the N-AS, is $K - plx - left - lceil \frac{M}{N} \rceil$. Particularly, the N-AS algorithm requires only the knowledge of the channel vector norms for selection. For this reason, the N-AS can be implemented without explicit channel estimation, supported by physical power-meters [21]. With this implementation, the N-AS requires a total of $2\tilde{}K$ symbols to operate. From this total, K symbols are required to estimate the norms of the channel vectors, and the remaining K symbols are used to estimate the channel vectors associated to the selected antennas.

TABLE II
Coordination Data Exchanged Between the RPUs and the CPU

Procedure	Implementation	Data type	Data size
GA-RA	Centralized	Channel matrix	MK
SCMAX-AS [14]	Centralized	Channel matrix	MK
N-AS [7]	Totally distributed	–	–
DGA-RA	Quasi-distributed	Gramian matrix	$(B + N_{\text{it}})K^2$

B. Coordination Data Size

The coordination data is defined as the data originated at the RPUs that is required at the CPU during the RA procedures. Determining the coordination data size is crucial since it can grow tremendously in the XL-MIMO scenario. In practical implementations, techniques as *data compression* helps alleviating the high interconnection bandwidth associated to the coordination data. However, such kind of consideration and optimization are out of the scope of this work.

Table II contains the coordination data size associated to the considered RA procedures, detailing the type of required data in each one. The GA-RA and SCMAX-AS procedures require the entire channel matrix at the CPU, while the DGA-RA one relies on the subarrays Gramian matrices. On the other hand, the N-AS procedure does not require any CSI knowledge at the CPU for antenna selection purpose, being the most appealing technique in terms of the coordination data size.

C. Number of Flops

The third complexity metric is the number of flops executed by each procedure. The complexity analysis for the N-AS and the GA-based AS algorithms are as follows. The SCMAX-AS procedure is not considered due to the high complexity associated with computing the number of executed operations by the convex optimization solver.

N-AS: The operations executed at each subarray on the N-AS procedure consists of calculating the channel vector norms then sorting the obtained values to get the N_b largest ones. Assuming that the sorting operation has the complexity of the order $M_b \log(M_b)$, the per-subarray flops for N-AS is

$$\mathcal{C}_{\text{N-AS}} = M_b(2\tilde{}K - 1) + M_b \log(M_b) \tag{25}$$

GA-RA: The complexity of the GA-RA method is dominated by the number of operations required for the evaluation of the GA

TABLE III
SIMULATION PARAMETERS

Parameter	Value
Cell size	$L = 30$ m
# Users	$K \in [1, 217]$
Maximum transmitted power	$P_{\max} = 230$ μW
Path-loss at the reference distance	$q_0 = -35.3$ dB
Path-loss exponent	$\kappa = 3$
Noise power	$\sigma_z^2 = -96$ dBm
Uniform Linear Array Setup	
# Antennas	$M \in [32, 2048]$
# RF transceivers	$N \in [64, 256]$
# Subarrays	$B = \{2, 4, 8\}$
# Antennas per subarray	$M_b = M/B$
# RF transceivers per subarray	$N_b = N/B$

fitness function, eq. (11). At the first iteration, the algorithm evaluate the fitness function for N_p individuals. During the remaining iterations, $(T-1)(N_p - N_e)$ fitness function evaluations are done, where T denotes the total number of generations.

As the OPA policy involves simple computations, the complexity of the fitness function is reduced to the inversion of the array Gramian matrix. The flops to compute the array Gramian matrix inverse is derived in Appendix XI. From this result, the total flops for the GA-RA algorithm is

$$\mathcal{C}_{\text{GA-RA}} = [T(N_p - N_e) + N_e] \left(\frac{7}{3}K^3 + 2NK^2 - K^2 \right) \quad (26)$$

DGA-RA: For the DGA-RA procedure, a similar approach to the one used for GA-RA can be followed. Despite that, the inverse of the array Gramian matrix is computed by the SMW formula, which is implemented with a different number of flops. The number of flops to obtain the inverse of the array Gramian matrix in the DGA-RA procedure is derived in Appendix XIII. Taking into account these differences and the fact that the DGA-RA procedure runs over N_{it} iterations, the total number of flops is given by:

$$\begin{aligned}\mathcal{C}_{\text{DGA-RA}} = N_{\text{it}} &[T(N_p - N_e) + N_e] \times \\ &\times \left[\frac{7}{3}N_b^3 + 2\tilde{}K^3 + N_b^2(4\tilde{}K - 1) + \right. \\ &\left. + K^2(4N_b - 2) + N_b^2(1 - 2\tilde{}K) + K \right] \quad (27)\end{aligned}$$

VI. NUMERICAL RESULTS

The numerical evaluations of the proposed methods as well as the benchmark techniques are presented in this section. The simulation system parameters are given in Table III. The users are randomly located inside a square cell of size L, and the BS is equipped with a uniform linear array (ULA) positioned on one side of the cell, as depicted in Fig. 1. Additionally, the users are random uniformly located at a distance in the range $(0.1\tilde{}L, L)$ from the array. Although the results in the following are obtained for the ULA, they can be easily extended to other array form factors, such as the uniform planar one.

Before comparing the proposed techniques, it is necessary to tune the GA-RA and DGA-RA GA input parameters in order to obtain a suitable performance-complexity trade-off. The input parameter N_p, p_c and p_m values are selected using the iterated local search algorithm [35]. The number of individuals for elitism is equal to 10% of the population size, and the number of tournaments is defined in order to fill the population after the elitism phase. Additionally, the stall convergence criterion parameter is approximately 30% of the maximum number of generations. The selected parameters for the GA-based procedures are listed in Table IV. Notice that the DGA-RA procedure is set to run 10 times less generations than the GA-RA, since the number of optimization variables decrease from M at the GA-RA to M_b in the DGA-RA procedure.

TABLE IV
GENETIC ALGORITHM PARAMETERS

Symbol	Description	Parameter value	
		GA-RA	DGA-RA
N_p	Population size	80	80
N_e	Elitism individuals	8	8
N_s	Tournaments	36	36
p_c	Crossover probability	0.33	0.35
p_m	Mutation probability	0.13	0.36
$T_{\max}$	Maximum generations	10^3	10^2
T_{stall}	Stall generations	300	30

In Fig. 4, the quality of convergence of the GA-RA procedure is corroborated varying the parameters N_p, p_c and p_m independently. Each surface is computed by averaging the achieved scores over 20 realizations. These results on the best and average SE scores among the generations t confirm the parameter values adopted in Table IV, while demonstrating a relative low tuning sensibility of the GA-RA convergence to the three input parameters.

Fig. 5 depicts the system SE achieved by the proposed RA procedures versus the number of available RF transceivers. In addition to the proposed solutions, the SE attained by random AS scheme and using all the M antennas are plotted as the lower and upper performance bounds, respectively. The results consider $M = 512$, $B = 8$, $K = 50$ and $N_{\text{it}} \in \{5, 16\}$ for the DGA-RA procedure. Observing the Fig. 5, one realize that the GA-based procedures achieve better SE results than the other ones. In the sequence, there are respectively the SCMAX-AS and N-AS. As expected, all the performance curves are upper and lower bounded by the SE achieved using full-array ZF and random AS, respectively. The SE gap between the procedures decreases as the number of RF transceivers increases. Analyzing the GA-based procedures, the DGA-RA achieves SE values tight to the GA-RA running with only five iterations. However, setting $N_{\text{it}} = 16$ makes the DGA-RA system SE values outperform marginally the ones obtained by the GA-RA procedure. Therefore, the quasi-distributed procedure can achieve a performance comparable, or even better, to the fully centralized approach by adopting a sufficient number of iterations.

In the following, Fig. 6 depicts the system SE achieved by the proposed RA procedures versus the number of users. These numerical results consider $M = 512$, $B = 8$, $N = 256$ and $N_{\text{it}} \in \{5, 16\}$ for the DGA-RA procedure. For better understanding, let $\mathcal{L} = K/N$ be the system effective loading factor. For all the proposed procedures, firstly the SE increases with K, assuming

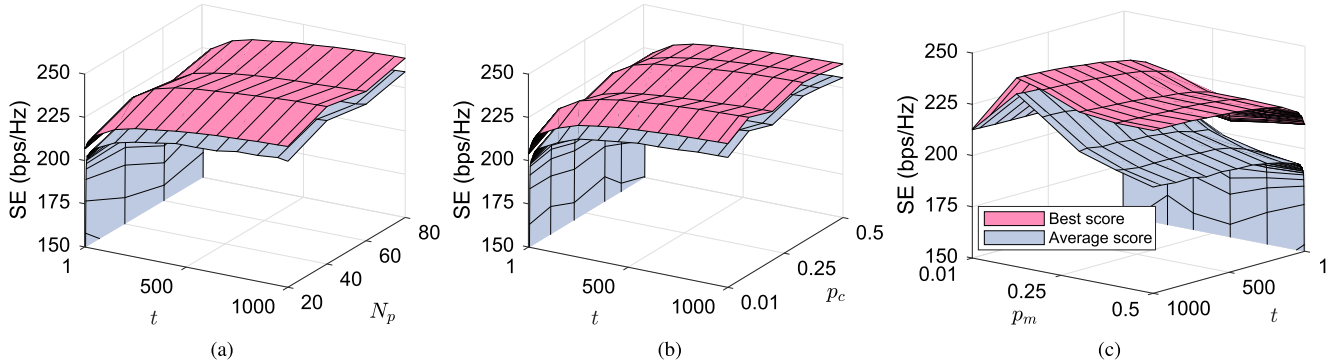


Fig. 4. Convergence of the GA-RA with the number of generations t varying the GA input parameters N_p, p_c and p_m. The "best" and "average" SE surfaces are obtained over 20 realizations. In each plot, the values of the remaining input parameters are given in Table IV.

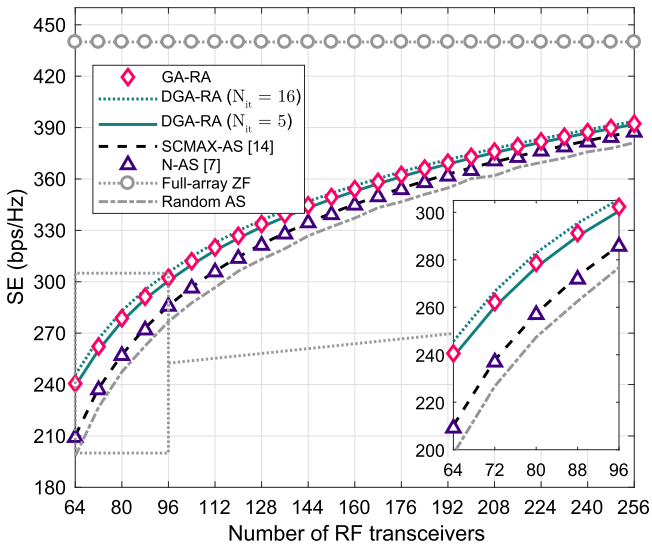


Fig. 5. Comparison of SE *vs* the number of available RF transceivers. $M = 512$, $B = 8$, $K = 50$ and, for the DGA-RA procedure $N_{\text{it}} \in \{5, 16\}$.

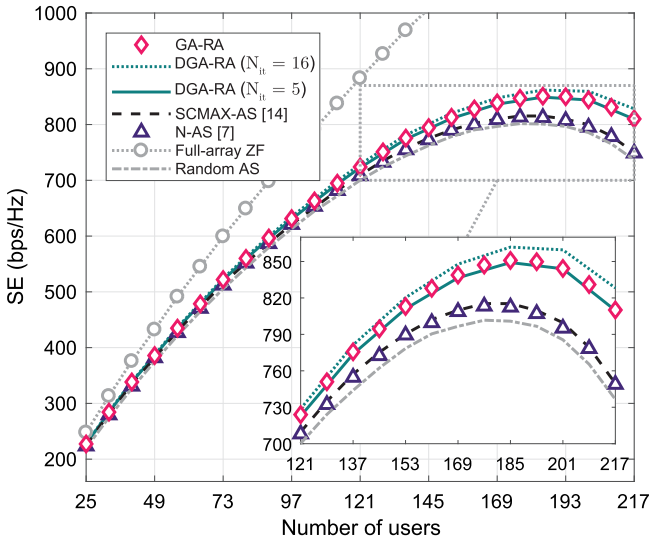


Fig. 6. Comparison of SE *vs* the number of users. $M = 512$, $B = 8$, $N = 256$ and, for the DGA-RA procedure $N_{\text{it}} \in \{5, 16\}$.

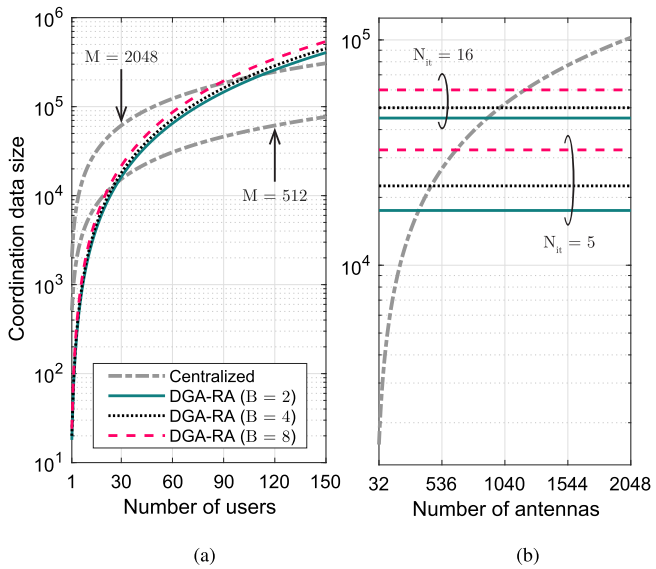


Fig. 7. Coordination data size of the GA-based RA schemes *vs* the number of (a) users and (b) antennas. When it is not specified, $N_{\text{it}} = 16$ and $K = 50$.

a decreasing behavior after a peak. This is due to the reduction of spatial degrees of freedom increasing the system loading factor, typically observed in linearly precoded systems [36]. Comparing the procedures, all of them get comparable SE values for a low loading factor. However, for high loading factor values, typically $\mathcal{L} = 0.6$, the GA-RA and DGA-RA procedures get substantial better results. Again, the DGA-RA outperforms the GA-RA in terms of SE by setting $N_{\text{it}} = 16$. Combining the results in Figs. 5 and 6, we conclude that the GA-based procedures perform with higher SE gains over the other available AS schemes [7], [14] in crowded XL-MIMO scenarios, *i.e.*, when the loading factor is high, $\mathcal{L} > 0.25$.

A. Complexity Analysis

The numerical results in the following cover the computational complexity of the proposed procedures. In Fig. 7(a) the coordination data size of the centralized procedures (GA-RA and SCMAX-AS) and the DGA-RA one versus the number of

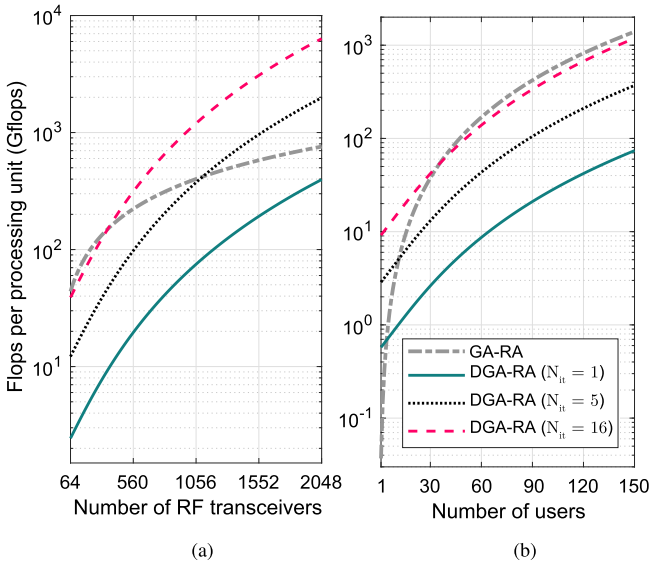


Fig. 8. Flops per processing unit of the proposed GA-based procedures *vs* the number of (a) available RF transceivers and (b) users. $B = 8$ and, when it is not specified, $K = 50$ and $N = 256$.

users is illustrated. The curves are evaluated by the expressions in Table II. The result considers $M \in \{512, 2048\}$ and, for the DGA-RA procedure, $N_{\text{it}} = 16$ and $B \in \{2, 4, 8\}$. Comparing the RA approaches when the number of users is low, the quasi-distributed one get lower coordination data sizes than the centralized procedures. For higher numbers of users, the coordination data size associated to DGA-RA acquires larger values than the obtained by the centralized procedures. This point of inversion of behavior depends on the numbers of antennas, subarrays and iterations w.r.t. the DGA-RA procedure. It is worth mentioning that the coordination data size grows quadratically with K for the DGA-RA procedure, while it grows linearly with K for the centralized RA procedure.

Fig. 7(b) depicts the coordination data size of the centralized procedures and the DGA-RA one versus the number of antennas in the BS. The results consider $K = 50$ and, for the DGA-RA method, $N_{\text{it}} \in \{5, 16\}$ and $B \in \{2, 4, 8\}$. The coordination data size grows linearly with M in the centralized procedures, while for the DGA-RA procedure, it does not depend on M. In fact, this is the primary aim for choosing a distributed RA technique in XL-MIMO, in which the BS is equipped with an asymptotically high number of antennas.

The next results are related to the complexity in terms of flops. Fig. 8(a) illustrates the number of flops per processing unit of the GA-based procedures versus the number of available RF transceivers. The curves are evaluated by the eqs. (26) and (27). Such results consider $K = 50$ and, for the DGA-RA procedure, $B = 8$ and $N_{\text{it}} \in \{1, 5, 16\}$. For low numbers of RF transceivers, the values of flops for the DGA-RA procedure are lower than the GA-RA algorithm. Again, after a point of inversion of behavior, the values of flops for GA-RA get lower than the ones for the quasi-distributed procedure. This point of changing of behavior decreases as N_{it} increases.

The curves with the number of flops per processing unit of the GA-based procedures versus the number of users are depicted in Fig. 8(b). This result considers $N = 256$ and, for the DGA-RA procedure, $B = 8$ and $N_{\text{it}} = \{1, 5, 16\}$. For low numbers of users, the values of flops of the GA-RA procedure are lower than the ones get for the DGA-RA. However, this behavior inverts quickly, and the gap between the values of flops for both centralized and distributed procedures becomes constant. This constant behavior for large K is due to the fact that both eqs. (26) and (27) grow asymptotically with K^3.

VII. Conclusion

This work proposes a subarray switching architecture for the BS antenna array, while examining the problem of joint AS and PA optimization aiming at maximizing the SE of XL-MIMO systems with limited number of RF transceivers. Two GA-based near-optimal and low-complexity procedures are proposed. One is the centralized GA-RA, designed to operate with the entire channel matrix available at the CPU. The other is the quasi-distributed DGA-RA, based on the subarrays Gramian matrices. Both evolutionary metaheuristic optimization methods are analyzed in terms of achieved SE, coordination data size and flops, and compared with benchmarks, including two procedures from the literature, the SCMAX-AS and the N-AS followed by optimal PA. Numerical results corroborate that the GA-based AS and PA procedures achieve high SE gains compared to the selected benchmarks, particularly in crowded XL-MIMO scenarios, *i.e.*, when the effective loading factor $\mathcal{L} > 0.25$. At the same time, the distributed DGA-RA method can outperform the other procedures with low-size coordination data and low computational complexity by taking the appropriate system operation settings.

Appendix A
Local Computation of the Inverse of the Array Gramian Matrix Via the Sherman-Morrison-Woodbury Formula

To compute the array Gramian matrix at the subarray b, the RPU must follow these two steps. Firstly, remove the contribution of the selected antennas at the subarray b at the iteration $n-1$. Then, add the contribution of the selected antennas at the iteration n. Therefore, it needs to compute the inverse of the array Gramian matrix by the expression

$$\left(\mathbf{G}_{\mathcal{S}}^{(n)}\right)^{-1} = \left(\mathbf{G}_{\mathcal{S}}^{(n-1)} - \mathbf{G}_{\mathcal{S}_b}^{(n-1)} + \mathbf{G}_{\mathcal{S}_b}^{(n)}\right)^{-1} \tag{28}$$

which evaluation would be straightforward if all the terms were available at the subarray.

However, the subarray needs to compute $(\mathbf{G}_{\mathcal{S}}^{(n)})^{-1}$ knowing only $(\mathbf{G}_{\mathcal{S}}^{(n-1)})^{-1}$ and the local channel vectors, *i.e.* $\underline{\mathbf{h}}_m \ \forall m \in \mathcal{M}_b$ for the subarray b. Writing the subarray Gramian matrices of (28) in terms of the local channel matrices results in

$$\begin{aligned} &- \mathbf{G}_{\mathcal{S}_b}^{(n-1)} + \mathbf{G}_{\mathcal{S}_b}^{(n)} \\ &\quad = - \left(\mathbf{H}_{\mathcal{S}_b}^{(n-1)}\right)^H \mathbf{H}_{\mathcal{S}_b}^{(n-1)} + \left(\mathbf{H}_{\mathcal{S}_b}^{(n)}\right)^H \mathbf{H}_{\mathcal{S}_b}^{(n)} \end{aligned}$$

TABLE V
FLOPS INVOLVED ON THE SHERMAN-MORRISON-WOODBURY FORMULA COMPUTATION

Symbol	Expression	Number of flops
$\mathbf{Q}_1$	$\mathbf{V}^H\mathbf{A}^{-1}$	$2N_bK^2 - N_bK$
$\mathbf{Q}_2$	$\mathbf{I} + \mathbf{Q}_1\mathbf{U}$	$2N_b^2K - N_b^2 + N_b$
$\mathbf{Q}_3$	$\mathbf{Q}_2^{-1}$	$7/3N_b^3$
$\mathbf{Q}_4$	$\mathbf{U}\mathbf{Q}_3$	$2N_b^2K - N_bK$
$\mathbf{Q}_5$	$\mathbf{I} - \mathbf{Q}_4\mathbf{Q}_1$	$2N_bK^2 - K^2 + K$
$\mathbf{Q}_6$	$\mathbf{A}^{-1}\mathbf{Q}_5$	$2K^3 - K^2$

$$= \left[-\left(\mathbf{H}_{\mathcal{S}_b}^{(n-1)}\right)^H \quad \left(\mathbf{H}_{\mathcal{S}_b}^{(n)}\right)^H \right] \begin{bmatrix} \mathbf{H}_{\mathcal{S}_b}^{(n-1)} \\ \mathbf{H}_{\mathcal{S}_b}^{(n)} \end{bmatrix} \tag{29}$$

From (28) and (29), it is possible to define the SMW formula variables, $\mathbf{A}^{-1}$, $\mathbf{U}$ and $\mathbf{V}^H$, in terms of the available information at the subarray as the eqs. (19), (20) and (21), respectively.

APPENDIX B
FLOPS TO COMPUTE THE INVERSE OF THE ARRAY GRAMIAN MATRIX VIA THE CHOLESKY DECOMPOSITION

Initially, the computation of the array Gramian matrix is done by solving the product in (5), which costs $2\tilde{}K^2\tilde{}N - K^2$ flops [34]. Afterwards, define the Cholesky decomposition of the array Gramian matrix as

$$\mathbf{G}_{\mathcal{S}} = \mathbf{L}\mathbf{L}^H \tag{30}$$

where $\mathbf{L}$ is a lower triangular matrix. The computation of $\mathbf{L}$ can be done with $K^3/3$ flops [34]. Then, each column of the inverse of the Gramian matrix can be computed solving the set of linear systems below by backforward substitution,

$$\mathbf{L}\mathbf{L}^H\mathbf{x} = \mathbf{e}_i, \ \forall i = 1, \ldots, K \tag{31}$$

where $\mathbf{e}_i$ denotes the canonical basis vector, *i.e.* a row vector with all entries equal to 0, except the entry i which is equal to 1. Each linear system can be solved with $2\tilde{}K^2$ flops [34], totaling $2\tilde{}K^3$ flops for all the columns of $\mathbf{G}_{\mathcal{S}}^{-1}$. Therefore, the total flops for the array Gramian matrix computation and inversion is equal to

$$\mathcal{C}_{Chol.} = \frac{7}{3}K^3 + 2NK^2 - K^2 \tag{32}$$

APPENDIX C
FLOPS TO COMPUTE THE INVERSE OF THE ARRAY GRAMIAN MATRIX VIA THE SHERMAN-MORRISON-WOODBURY FORMULA

To count the flops to compute the matrix inversion by the SMW formula, the eq. (18) is decomposed in six parts. The computations involved in each part and their respective flops are organized in Table V. The flops in Table V are counted assuming that the contribution of the selected antennas during the previous iteration is removed. Such assumption is reasonable since the expression in (28) can be done sequentially, by keeping only the terms $-\mathbf{G}_{\mathcal{S}_b}^{(n-1)}$ or $\mathbf{G}_{\mathcal{S}_b}^{(n)}$ at a time.

All the parts include only simple matrix multiplications and sums, except for the part $\mathbf{Q}_3$. This part can be efficiently computed by the Cholesky decomposition approach followed by the backforward substitution procedure described in Appendix XI. Therefore, the total flops required to compute the inverse of the array Gramian matrix via the SMW formula is equal to

$$\mathcal{C}_{SMW} = \frac{7}{3}N_b^3 + 2\tilde{}K^3 + N_b^2(4\tilde{}K - 1) + K^2(4N_b - 2) + N_b^2(1 - 2\tilde{}K) + K \tag{33}$$

REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[2] E. D. Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath, "Non-stationarities in extra-large-scale massive MIMO," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 74–80, Aug. 2020.

[3] Á. O. Martínez, E. De Carvalho, and J. Ã. Nielsen, "Towards very large aperture massive MIMO: A measurement based study," in *Proc. IEEE Globecom Workshops*, Dec. 8-12 2014, pp. 281–286.

[4] Z. Zhou, X. Gao, J. Fang, and Z. Chen, "Spherical wave channel and analysis for large linear array in LoS conditions," in *Proc. IEEE Globecom Workshops*, Dec. 6-10 2015, pp. 1–6.

[5] X. Li, S. Zhou, E. Björnson, and J. Wang, "Capacity analysis for spatially non-wide sense stationary uplink massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 7044–7056, Dec. 2015.

[6] A. Ali, E. D. Carvalho, and R. W. Heath, "Linear receivers in non-stationary massive MIMO channels with visibility regions," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 885–888, Jun. 2019.

[7] A. Garcia-Rodriguez, C. Masouros, and P. Rulikowski, "Reduced switching connectivity for large scale antenna selection," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2250–2263, May 2017.

[8] Y. Gao, H. Vinck, and T. Kaiser, "Massive MIMO antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1346–1360, Mar. 2018.

[9] K. Li, R. R. Sharan, Y. Chen, T. Goldstein, J. R. Cavallaro, and C. Studer, "Decentralized baseband processing for massive MU-MIMO systems," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 491–507, Dec. 2017.

[10] J. R. Sánchez, F. Rusek, O. Edfors, M. Sarajlić, and L. Liu, "Decentralized massive MIMO processing exploring daisy-chain architecture and recursive algorithms," *IEEE Trans. Signal Process.*, vol. 68, pp. 687–700, Jan. 2020.

[11] A. Mueller, A. Kammoun, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Reducing complexity in massive MIMO," *EURASIP J. Wireless Commun. Netw.*, no. 63, pp. 1687–1499, Feb. 2016.

[12] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[13] A. Dua, K. Medepalli, and A. J. Paulraj, "Receive antenna selection in MIMO systems using convex optimization," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2353–2357, Sep. 2006.

[14] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[15] P. Lin and S. Tsai, "Performance analysis and algorithm designs for transmit antenna selection in linearly precoded multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1698–1708, May 2012.

[16] A. Amiri, C. N. Manchon, and E. de Carvalho, "Deep learning based spatial user mapping on extra large MIMO arrays," Feb. 2020, *arXiv. 2002.00474*.

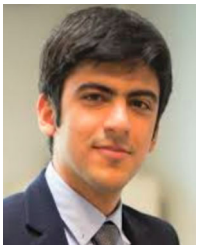
[17] J. C. Marinello, T. Abrão, A. Amiri, E. de Carvalho, and P. Popovski, "Antenna selection for improving energy efficiency in XL-MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 305–13 318, Nov. 2020.

[18] H. Lu and W. Fang, "Joint transmit/receive antenna selection in MIMO systems based on the priority-based genetic algorithm," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 588–591, Dec. 2007.

[19] J. Lain, "Joint transmit/receive antenna selection for MIMO systems: A real-valued genetic approach," *IEEE Commun. Lett.*, vol. 15, no. 1, pp. 58–60, Jan. 2011.

[20] B. Makki, A. Ide, T. Svensson, T. Eriksson, and M. Alouini, "A genetic algorithm-based antenna selection approach for large-but-finite MIMO networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 6591–6595, Jul. 2017.
[21] J. R. Sánchez, J. Vidal Alegría, and F. Rusek, "Decentralized massive MIMO systems: Is there anything to be discussed?," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 7-12 2019, pp. 787–791.
[22] A. Amiri, M. Angjelichinoski, E. de Carvalho, and R. W. Heath, "Extremely large aperture massive MIMO: Low complexity receiver architectures," in *Proc. IEEE Globecom Workshops*, Dec. 9-13 2018, pp. 1–6.
[23] A. Amiri, C. N. Manchón, and E. de Carvalho, "A message passing based receiver for extra-large scale MIMO," in *Proc. IEEE 8th Int. Workshop Comput. Adv. Multi-Sensor Adaptive Process.* (CAMSAP), 2019, pp. 564–568.
[24] A. Amiri, S. Rezaie, C. N. Manchon, and E. de Carvalho, "Distributed receivers for extra-large scale MIMO arrays: A message passing approach," Jul. 2020, *arXiv:2007.06930*.
[25] X. Yang, F. Cao, M. Matthaiou, and S. Jin, "On the uplink transmission of multi-user extra-large scale massive MIMO systems," 2019, *arXiv:1909.06760*.
[26] D. A. Gore, R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, vol. 5, Jun. 5-9 2000, pp. 2785–2788.
[27] A. F. Molisch, M. Z. Win, Yang-Seok Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1759–1772, Jul. 2005.
[28] S. Asaad, A. M. Rabiei, and R. R. Müller, "Massive MIMO with antenna selection: Fundamental limits and applications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8502–8516, Dec. 2018.
[29] C. Ouyang, Z. Ou, L. Zhang, P. Yang, and H. Yang, "Asymptotic upper capacity bound for receive antenna selection in massive MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, May. 20–24 2019, pp. 1–6.
[30] Z. Abdullah, C. C. Tsimenidis, G. Chen, M. Johnston, and J. A. Chambers, "Efficient low-complexity antenna selection algorithms in multi-user massive MIMO systems with matched filter precoding," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2993–3007, Mar. 2020.
[31] H. Siljak, I. Macaluso, and N. Marchetti, "Distributing complexity: A new approach to antenna selection for distributed massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 902–905, Dec. 2018.
[32] P. He, L. Zhao, S. Zhou, and Z. Niu, "Water-filling: A geometric approach and its application to solve generalized radio resource allocation problems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3637–3647, Jun. 2013.
[33] S. Luke, *Essentials of Metaheuristics*. Morrisville, NC, USA: Lulu, 2013.
[34] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2013.
[35] E. Montero, M.-C. Riff, and B. Neveu, "A beginner's guide to tuning methods," *Appl. Soft Comput.*, vol. 17, pp. 39–51, Apr. 2014.
[36] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, Cambridgeshire, U.K.: Cambridge Univ. Press, 2016.

João Henrique Inacio de Souza (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from the State University of Londrina (UEL), Londrina, Brazil, in 2018. He is currently working toward the Ph.D. degree in electrical engineering with the UEL/UTFPR Program supported by the National Council for Scientific and Technologic Development (CNPq). His research interests include optimization of wireless communications systems aspects, cross-layer resource allocation, massive multi-antenna systems, and access protocols.



Abolfazl Amiri received the B.S. degree in electrical and communications engineering (with honors) from the University of Tabriz, Tabriz, Iran, in 2013 and the M.S. degree in electrical and communication systems engineering from the University of Tehran, Tehran, Iran, in 2016. He is currently a Ph.D. Fellow with Electronic Systems Department, Aalborg University, Aalborg, Denmark. From 2016 to 2017, he was a RF Engineer and Cellular Network Optimizer with Huawei technologies Tehran office. His research interests include applications of signal processing and machine learning in wireless communications and multiantenna systems.

Taufik Abrão (Senior Member, IEEE) received the B.S., M.Sc., and Ph.D. degrees in electrical engineering from the Polytechnic School of the University of São Paulo, São Paulo, Brazil, in 1992, 1996, and 2001, respectively. Since March 1997, he has been with the Communications Group, Department of Electrical Engineering, Londrina State University, Paraná, Brazil, where he is currently an Associate Professor of telecommunications and the Head of the Telecomm. & Signal Processing Lab. In 2018, he was with the Connectivity section, Aalborg University, Aalborg, Denmark, as a Guest Researcher. In 2012, he was an Academic Visitor with the Southampton Wireless Research Group, University of Southampton, Southampton, U.K. From 2007 to 2008, he was a Postdoctoral Researcher with the Polytechnic University of Catalonia, Barcelona, Spain. His current research interests include communications and signal processing, especially massive MIMO, XL-MIMO, URLLC, mMTC, optimization methods, machine learning, detection, estimation, resource allocation, and protocols. He was an Associate Editor for the IEEE SYSTEMS JOURNAL in 2020, the IEEE ACCESS during 2016–2018, IEEE COMMUNICATIONS SURVEYS AND TUTORIALS during 2013–2017, the AEUe-Elsevier in 2020, the *IET Signal Processing* in 2019, and JCIS-SBrT during 2018–2020, and an Executive Editor of the ETT-Wiley journal during 2016–2021.

Elisabeth de Carvalho (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Telecom ParisTech, Paris, France. She was a Postdoctoral Fellow with Stanford University, Standford, CA, USA, and then with industry in the field of DSL and wireless LAN. Since 2005, she has been an Associate Professor with Aalborg University, Aalborg, Denmark, where she has led several research projects in wireless communications. She has coauthored the text book *A Practical Guide to the MIMO Radio Channel*. Her main expertise is in signal processing for MIMO communications, with recent focus on massive MIMO, including channel measurements, channel modeling, beamforming, and protocol aspects.

Petar Popovski (Fellow, IEEE) received the Dipl.-Ing and M.Sc. degrees in communication engineering from the University of Sts. Cyril and Methodius in Skopje, and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 2005. He is currently a Professor with Aalborg University, where he heads the section on Connectivity and a Visiting Excellence Chair with the University of Bremen, Bremen, Germany. He has authored the book *Wireless Connectivity: An Intuitive and Fundamental Guide*, published by Wiley in 2020. His research interests include wireless communication and communication theory. He was the recipient of the ERC Consolidator Grant (2015), the Danish Elite Researcher Award (2016), IEEE Fred W. Ellersick Prize (2016), IEEE Stephen O. Rice Prize (2018), Technical Achievement Award from the IEEE Technical Committee on Smart Grid Communications (2019), the Danish Telecommunication Prize (2020), and Villum Investigator Grant (2021). He is a Member at Large at the Board of Governors in IEEE Communication Society, the Vice-Chair of the IEEE Communication Theory Technical Committee and IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. He is currently an Area Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was the General Chair for IEEE SmartGridComm 2018 and IEEE Communication Theory Workshop 2019.

APPENDIX B – Deep Learning-based Activity Detection for Grant-free Random Access

Deep Learning-based Activity Detection for Grant-free Random Access

João Henrique Inacio de Souza, Taufik Abrão

***Abstract*—The cellular internet-of-things (IoT) wireless network is a promising topology to provide massive connectivity for machine-type devices. However, designing grant-free random access (GF-RA) protocols to manage such connections is challenging, since they must operate in interference-aware scenarios with sporadic device activation patterns and shortage of mutually orthogonal resources. Supervised machine learning models have been provided efficient solutions for non-orthogonal preamble design, activity detection, as well as non-coherent data detection in scenarios with massive connectivity. Considering these promising results, in this work we propose two deep learning (DL) sparse support recovery algorithms for activity detection in random access (RA) communications, one based on the deep feedforward network model, and another based on the convolutional neural network model aiming to deploy GF-RA protocols. Differently from other works, we investigate the impact of the type of sequences for preamble design on the activity detection performance. Our results reveal that preambles based on the Zadoff-Chu sequences, which present good cross-correlation properties, achieve better performance with the proposed algorithms. Besides, we demonstrate that our DL algorithms achieve performance comparable to state-of-art techniques with extremely-low computational complexity.**

***Index Terms*—Massive machine-type communications (mMTC), random access protocols, grant-free access, deep feedforward networks, convolutional neural networks.**

I. Introduction

The massive machine-type communication (mMTC) service will provide full connectivity for internet-of-thing (IoT) applications in cellular wireless networks. The machine-type devices have sporadic activation patterns, transmit small payloads with low-complexity network hardware and are powered by short lifetime batteries. At the same time, they appear as increasing massive populations distributed over wide areas. All these characteristics combined impose challenges to provide connectivity, principally in regarding to network access. The limited channel coherence time and the high number of devices prohibit the design of mutually-orthogonal preamble sequences. Operating with fewer preambles than devices result in resource contention and network collision, which increases the access delay due to retransmissions. On the other hand, non-orthogonal preambles reduce the number of collisions, at the cost of introducing interference among the active devices and degrading the system performance [1].

This work was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001 and in part by the National Council for Scientific and Technological Development (CNPq) of Brazil under Grant 310681/2019-7.

J. H. I. de Souza and T. Abrão are with the Electrical Engineering Department, State University of Londrina, PR, Brazil. E-mail: joaohis@outlook.com; taufik@uel.br.

The grant-free random access (GF-RA) is an alternative to provide connectivity dealing efficiently with the interference generated by the non-orthogonal preambles, while keeping a low network access delay [2].

Methods for *sparse signal processing* are broadly employed to cope with the interference and enable GF-RA procedures. These techniques are suitable for the typical mMTC scenario due to the sparse nature of the signals generated by machine-type devices. Specifically, we are interested in the sparse support and signal recovery methods. *Sparse support recovery* aims to estimate the indices of the active devices during a random access slot transmission. In contrast, *sparse signal recovery* is employed for channel estimation or data detection. Approaches to perform sparse support and signal recovery are classified in compressed sensing (CS)-based, covariance (CV)-based [2] and, more recently, machine learning (ML)-based.

The sparse support and signal recovery based on ML approaches have presented results which outperform substantially the other two CS-based and CV-based techniques, both in performance and computational complexity. These methods have extremely-low run time at the cost of an intense off-line training phase to learn the neural network weights. Furthermore, the *universal approximation theorem* states that properly-designed neural networks are capable of approximate any Borel measurable function with an arbitrary non-zero amount of error [3].

A. Literature Review

We provide an outline of the literature on sparse support and signal recovery applicable to GF-RA protocols. The approximate message passing (AMP) algorithm from the CS literature is present in many recent contributions for activity detection and channel estimation purpose [4], [5]. Specifically, the authors in [5] propose a non-coherent scheme that embed information bits into the devices preamble sequences, allowing joint activity and data detection. Despite the positive results involving the AMP algorithm, its dependence on statistical channel information represents a technical difficulty for practical implementations.

Iterative methods using the empirical covariance matrix of the received signal are used for activity detection in [6]. The authors study two low-complexity algorithms for activity detection in massive MIMO unsourced random access. Despite the scheme being non-coherent, the computational complexity is still high owing to the iterative nature of the detection algorithms.

When it comes to ML methods, we can separate them as *model-driven* and *data-driven* approaches. The model-driven

approaches are networks designed to mimic the detection process of existing sparse support recovery algorithms. On the other hand, data-driven approaches are conventional network architectures trained to learn the inverse mapping from noisy received symbols to the original transmitted signals. Model-driven approaches obtained by unfolding the AMP algorithm and variations are given in [7], [8]. Additionally, the authors in [9] parameterize the iterative shrinkage thresholding algorithm (ISTA), implementing a recurrent neural network (RNN) to perform joint activity detection and channel estimation. The devised RNN outperforms the classic ISTA. On the matter of the data-driven approaches, we have the network of [10] which performs active device detection in non-orthogonal multiple access (NOMA), while the scheme of [11] is capable of adjust dynamically its number of layers for sparse signal recovery. Specially, in [12] a deep auto-encoder for jointly design the preamble matrix and the activity detection algorithm is developed. The proposed approach achieves a much lower error rate than classic methods with lower computational complexity, due to the efficient joint design. The results are extended to the massive MIMO receiver in [13].

The authors in [14] developed a framework to recover sparse signals from noisy measurements in imaging applications using the convolutional neural network (CNN). The scheme learns both the representation for signals and the inverse mapping function from measurement vectors. Compared with the conventional densely connected networks, the convolutional layers have as advantage the sparse connections between the neurons and parameter sharing. The scheme in [14] presents a good tradeoff between the reconstruction time and probability of successful signal recovery, approximating the solution of state-of-art algorithms in a much lower run time.

B. Contributions

We develop two deep learning (DL) algorithms for *sparse support recovery*, used for efficient activity detection in machine-type GF-RA protocols. The networks are subjected to an off-line training with the backpropagation algorithm and a representative data set to learn the weights in order to obtain accurate activity detection. The algorithms, which inputs are the received symbols during the random access slot, are designed using a deep feedforward network (DFN) with densely connected layers, and also a CNN built on stacked convolutional layers.

Contributions: To our knowledge, it is the first time that a CNN is used for activity detection of machine-type communication devices. Besides, unlike works in the recent literature that use model-driven approaches [7]–[9], we use a data-driven approach. We extensively describe the architecture of the networks, presenting the procedures for parameter tuning and network training. Then, we conduct experiments to evaluate the performance of the activity detection algorithms in terms of error rate and computational complexity. Differently from [12], we focus on a broad analysis of the types of error occurred during activity detection, as they have different impacts on the system performance according to the application. Additionally, we evaluate the algorithms with three kinds of preamble sequences: Gaussian and Bernoulli random sequences, as well as Zadoff-Chu (ZC) sequences [15], which present good cross-correlation properties. The numerical results reveal that the DL algorithms achieve accurate activity detection with extremely-low computational run time owing to the networks architecture. Specifically, our CNN-based activity detection scheme has its performance improved with the ZC, thanks to the good cross-correlation properties of ZC sequences. The extensive numerical results corroborate both the performance gains and improvements on the computational complexity attained by the proposed algorithms when compared with the reference methods.

II. System Model

In this section, we describe the system model, as well as devise the activity detection problem. We consider the uplink of a narrow-band system constituted by one cell with K machine-type devices served by a single-antenna base-station, as Fig. 1 depicts. These devices access the network with probability $p_a \ll 1$. For this reason, the number of active devices is much less than the total in the cell.

During the random access slot, the active devices transmit unique preamble sequences. These sequences identify the devices and are used to performing activity detection at the base-station. The preamble sequences have a length of $L < K$ symbols due to the limited channel coherence time. For this reason, it is impossible to design mutually-orthogonal sets of preambles. In our work, we consider three types of non-orthogonal sequences for the preambles. The non-orthogonal sequences result in interference on the transmitted signal by the set of active devices during a random access slot. Details on the used sequences are given in the next subsection. The preamble sequences are defined by the vectors $\mathbf{a}_k \in \mathbb{C}^{L\times 1}$ for $k = 1, \ldots, K$, and are normalized such that $\|\mathbf{a}_k\|_2^2 = L, \forall k$. The preamble matrix $\mathbf{A} \in \mathbb{C}^{L\times K}$, which is known at the base-station, contains in its columns the preamble sequences of all the K devices in the cell.

Let $\alpha_k \in \{0, 1\}$ for $k = 1, \ldots, K$ be an activity state indicator of the device k, which is equal to 1 if the device is active and 0 otherwise. We consider that the probability

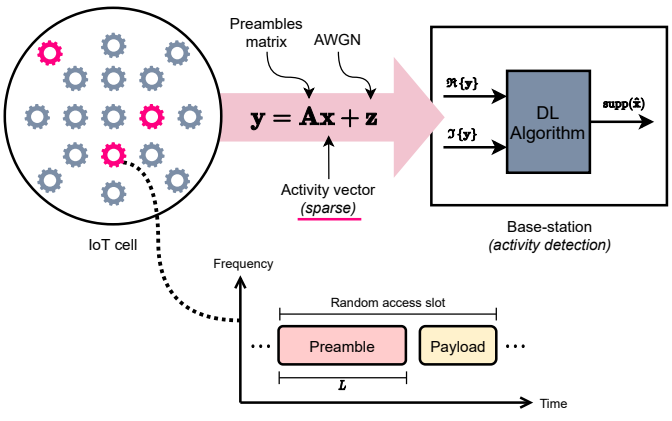


Figure 1. System diagram with an activity detection scheme based on a deep learning algorithm and the random access slot.

of each device being active follows a Bernoulli distribution, *i.e.* $\alpha_k \sim p_{\alpha_k}(u) = (1-p_a)\delta(u) + p_a\delta(u-1), \forall k$, where $p_{\alpha_k}(u)$ is the probability mass function (PMF) of α_k and $\delta(u)$ is the Dirac delta function. The vector with the activity state indicators of all the K devices is named **activity descriptor**,

$$\boldsymbol{\alpha} = \begin{bmatrix}\alpha_1 & \cdots & \alpha_K\end{bmatrix}^T \tag{1}$$

The channel links between each device and the base-station follow the Rayleigh fading model, with channel coefficients associated to the device k defined by $h_k \sim \mathcal{CN}(0,1)$.

The received signal at the base-station during random access slot is equal to the superposition of the preambles transmitted by each active device, which is written as

$$\mathbf{y} = \sum_{k=1}^{K} \alpha_k h_k \mathbf{a}_k + \mathbf{z} \tag{2}$$

where $\mathbf{z} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_z^2\mathbf{I}\right)$ is the additive white Gaussian noise. Letting $\mathbf{x} = [\alpha_1 h_1 \ \cdots \ \alpha_K h_K]^T$ allows us to write the compact form of the received signal

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z} \tag{3}$$

The goal of the **activity detection problem** is to determine the set of active devices from the received signal $\mathbf{y}$ and the preamble matrix $\mathbf{A}$. This is equivalent to finding the index of the non-zero entries of the vector $\mathbf{x}$, namely the support of $\mathbf{x}$,

$$\operatorname{supp}(\mathbf{x}) = \{k \in \mathbb{N}_+ \,|\, x_k \neq 0\} \tag{4}$$

As the number of active devices is much less than the total number of devices in the cell (sporadic activation, *i.e.* $p_a \ll 1$), $\mathbf{x}$ has a sparse structure. For this reason, the problem of finding the support of $\mathbf{x}$ from $\mathbf{y}$ and $\mathbf{A}$ can be cast as a *sparse support recovery* problem.

The *sparse support recovery* problem is an object of study of the compressed sensing field, and have been addressed in many works using different approaches, *e.g.* see [16] and the references therein. In our work, the aim is to develop two DL-based algorithms for sparse support recovery in the context of activity detection for massive machine-type communications. We demonstrate that a DL algorithm which inputs are the real and imaginary parts of the received signal is sufficient to perform activity detection efficiently.

A. Non-orthogonal Sequences

The preamble sequences play a key role in the activity detection problem, as they dictates the interference levels on the transmitted signal by the active devices during a random access slot. In the context of sparse support recovery using CS techniques, it is common to use preambles generated by sampling random distributions. However, we show that using deterministic sequences with good correlation properties in a DL technique achieves good activity detection results. In the following, we present suitable types of sequences used in our work.

Random Sequences. We use two types of random sequences for the preambles. Random sequences are useful, specially in cases where the number of mutually orthogonal preambles is not sufficient for all devices, *e.g.* in crowded mMTC applications, because it is easy to generate large amount of unique preambles. Additionally, matrices generated by sampling sub-Gaussian random distributions (*e.g.* Normal and Bernoulli) satisfy, with high probability, the restricted isometry property (RIP), a necessary condition to guarantee the signal reconstruction of many CS algorithms [17]. The first type of random sequence is obtained by sampling a circularly-symmetric complex Normal distribution with zero mean and unitary variance:

$$[\bar{\mathbf{a}}_k^{\mathrm{G}}]_l \sim \mathcal{CN}(0,1) \tag{5}$$

for $k = 1, \ldots, K$ and $l = 1, \ldots, L$. We name this type of sequence as Normal. In order to met the preamble normalization presented previously, we scale the Normal preamble with the factor:

$$\mathbf{a}_k^{\mathrm{G}} = \sqrt{\frac{L}{\|\bar{\mathbf{a}}_k^{\mathrm{G}}\|_2^2}}\bar{\mathbf{a}}_k^{\mathrm{G}} \tag{6}$$

The second type of random sequence type is the Bernoulli. The Bernoulli sequences are real two-valued sequences obtained by sampling a symmetric random distribution of the type

$$[\mathbf{a}_k^{\mathrm{B}}]_l \sim p_{\mathrm{B}}(u) = \frac{1}{2}\delta(u+1) + \frac{1}{2}\delta(u-1) \tag{7}$$

for $k = 1, \ldots, K$ and $l = 1, \ldots, L$. As each entry of the Bernoulli sequence has norm equal to 1, normalization is not required.

Zadoff-Chu Sequences. The ZC sequence is a type of polyphase sequence which entries are defined by [15]:

$$[\bar{\mathbf{a}}_r^{\mathrm{ZC}}]_l = \exp\left[-\frac{j\pi r}{L}(l-1)l\right], \, l = 1, \ldots, L \tag{8}$$

where $j = \sqrt{-1}$ and $r \in \{1, \ldots, L-1\}$ is a number relatively prime to L named the **sequence root**. A ZC sequence has the ideal auto-correlation property, *i.e.* its auto-correlation value is equal to zero for all shifts other than zero. For this reason, a sequence and its shifted versions comprise a set of mutually-orthogonal sequences. The ideal auto-correlation property holds only for sequences generated by a single root. On the other hand, sequences generated by different roots have constant cross-correlation equal to $\sqrt{L}$ if the difference between the roots is relatively prime to L [18]. Therefore, a set of non-orthogonal sequences with a three-valued cross-correlation function can be generated by taking the shifted versions of multi-root ZC sequences. We use this set of non-orthogonal sequences to generate the ZC preambles. The ZC preambles are defined by the sequences

$$\bar{\mathbf{a}}_r^{\mathrm{ZC}}, \forall r \in \mathcal{R} \tag{9}$$

and its shifted versions, considering $\mathcal{R} \subset \{1, \ldots, L-1\}$ the set of chosen roots. Given the number of devices in the cell and the preamble length, the *minimum number of roots to generate unique preambles* for all the K devices is equal to

$$N_r = \left\lceil \frac{K}{L} \right\rceil \tag{10}$$

It is worth mentioning that the set of non-orthogonal ZC preambles is composed by $|\mathcal{R}| \geq N_r$ smaller sub-sets of orthogonal ones. For this reason, allocating the sub-set of orthogonal preambles to devices with similar activation pattern is an efficient alternative to manage the interference levels.

B. Sequences Performance in Activity Detection Problem

In order to give an insight of the impact of the preamble design on the system performance, we demonstrate a result on the distribution of the signal-to-interference-plus-noise (SINR) ratio of the received signal at the base-station for the analyzed sequence types. The SINR gives an indirect measure of activity detection performance, as it demonstrates the ratio between the power of the signal of an active device and the power of noise plus interference generated by the simultaneous interfering devices. Fig. 2 depicts the empirical cumulative distribution function of the SINR for the k-th device by using the Normal, Bernoulli and ZC sequences, considering three values of activation probability, $p_a \in \{0.01, 0.10, 0.30\}$. The remain setup parameter values to generate this result are $K = 40$, $L = 20$ and $\sigma_z^2 = 0.1$, resulting in a signal-to-noise ratio (SNR) of 10 dB. Besides, we define the SINR correlating the received signal at the BS with the preamble of each active device, resulting:

$$\text{SINR}_k = \frac{\|\mathbf{a}_k\|_2^4 |h_k|^2}{\sum_{i \in \mathcal{K}_\text{A}, i \neq k} |\mathbf{a}_k^H \mathbf{a}_i|^2 |h_i|^2 + |\mathbf{a}_k^H \mathbf{z}|^2} \tag{11}$$

for $k \in \mathcal{K}_\text{A}$, where $\mathcal{K}_\text{A}$ is the set of active users during a specific random access slot. From Fig. 2, one can observe in the set of curves that decreasing the activation probability increases the average SINR values, since it reduces the number of simultaneous interfering devices during a specific transmission interval. The ZC sequences achieve the best average SINR values due to their good (reduced) cross-correlation properties. At the same time, the Bernoulli sequences outperform marginally the Normal ones in terms of the SINR distribution.

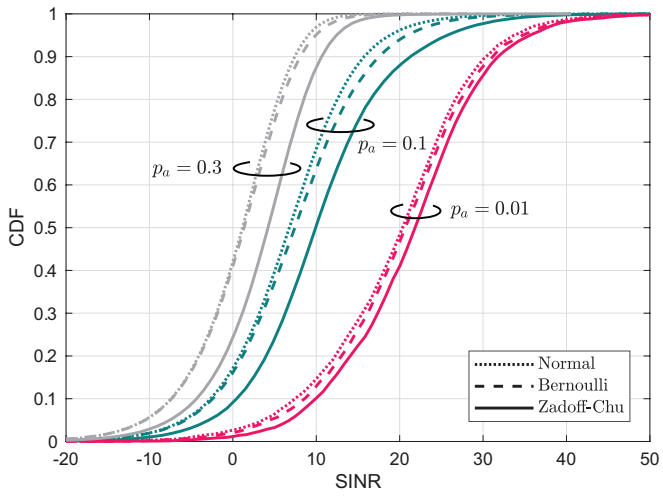


Figure 2. Empirical CDF of the SINR using the Normal, Bernoulli and Zadoff-Chu sequences for three values of p_a.

Table I
DESCRIPTION OF THE LAYERS IN THE DFN ALGORITHM.

Layer	Type	Activation function	Neurons
Hidden 1	Densely connected	ReLU	N_n
Hidden 2	Densely connected	ReLU	N_n
Output	Densely connected	Sigmoid	K

III. Deep Learning Algorithms

In this section, we introduce the architectures of the DFN and CNN algorithms devised for activity detection. The DL algorithms are represented by functions of the type $f(\mathbf{v}_0, \theta) : \mathcal{I} \rightarrow \mathbb{R}^{K \times 1}$, where $\mathbf{v}_0 \in \mathcal{I}$ are the network inputs in the input space $\mathcal{I}$ and θ is a set with the network parameters, which learns an inverse map of the activity descriptor from the received signal during the random access slot. The learning phase is implemented by a training procedure, in which the parameters $\boldsymbol{\theta}$ of the function $f(\mathbf{v}_0, \boldsymbol{\theta})$ are adjusted following a set of examples containing the activity descriptor and the respective received signal at the base-station.

Let $\left\{\left(\mathbf{y}^{(s)}, \boldsymbol{\alpha}^{(s)}\right) | s = 1, \ldots, S\right\}$ be a set of samples containing the received signal and the activity descriptor of a random access slot. We denote the output of a DL algorithm for the sample s as $\tilde{\boldsymbol{\alpha}}^{(s)}$. Taking this into account, we can measure the quality of the estimate $\tilde{\boldsymbol{\alpha}}^{(s)}$ with respect to the original activity descriptor $\boldsymbol{\alpha}^{(s)}$ by calculating the binary cross-entropy function. In the DL methods, the loss function is defined as the *average binary cross-entropy* function [19], calculated over all the S samples as:

$$\mathcal{L} = \frac{1}{SK} \sum_{s=1}^{S} \sum_{k=1}^{K} - \left[\alpha_k^{(s)} \log\left(\tilde{\alpha}_k^{(s)}\right) + \left(1 - \alpha_k^{(s)}\right) \log\left(1 - \tilde{\alpha}_k^{(s)}\right)\right] \tag{12}$$

The aim of the training procedure is to minimize $\mathcal{L}$ by choosing the right parameters of $f(\mathbf{v}_0, \boldsymbol{\theta})$. In this way, the trained DL network produces an accurate estimate of the original activity descriptor.

A. Deep Feedforward Network

The DFN algorithm, as depicts the Fig. 3, is implemented with densely connected layers. Its inputs are the real and imaginary parts of the received signal, while the output is an estimate of the activity descriptor. In the following, we describe in detail the layers of the DFN.

The DFN architecture has two densely connected hidden layers with N_n neurons and one densely connected output layer with K neurons. The hidden layers have the ReLU activation function:

$$\text{ReLU}(u) = \max(0, u) \tag{13}$$

On the other hand, the output layer has the sigmoid activation function:

$$\text{Sigmoid}(u) = \frac{1}{1 + \exp(-u)} \tag{14}$$

The details of the DFN layers are organized in Table I.

Now, we introduce the parameters of the DFN layers, which are the weights matrices and the bias values vectors. Given the

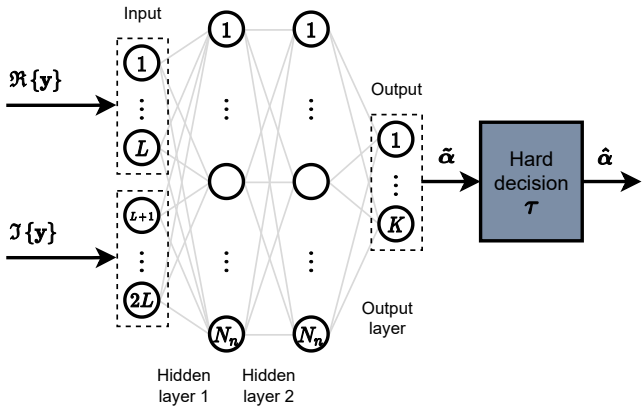


Figure 3. Architecture of the DFN algorithm proposed for activity detection.

dimensions of the inputs and outputs of each layer, we have the weights matrices $\mathbf{W}_1 \in \mathbb{R}^{N_n \times 2L}$, $\mathbf{W}_2 \in \mathbb{R}^{N_n \times N_n}$ and $\mathbf{W}_3 \in \mathbb{R}^{K \times N_n}$ for the hidden layers 1 and 2, and the output layer, respectively. Similarly, the vectors with the bias values are $\mathbf{b}_1 \in \mathbb{R}^{N_n \times 1}$, $\mathbf{b}_2 \in \mathbb{R}^{N_n \times 1}$ and $\mathbf{b}_3 \in \mathbb{R}^{K \times 1}$. Taking this into account, we define the set with the DFN parameters by

$$\boldsymbol{\theta}_{\text{DFN}} = \{\mathbf{W}_i, \mathbf{b}_i\}_{i=1}^{3} \tag{15}$$

Considering the DFN parameters, we write the input-output relation of the hidden layers in the DFN:

$$\mathbf{v}_i = \text{ReLU}\left(\mathbf{W}_i \mathbf{v}_{i-1} + \mathbf{b}_i\right), \; i = 1, 2 \tag{16}$$

where $\mathbf{v}_i$ is the layer output, considering $\mathbf{v}_0$ as the DFN input, and the activation function is calculated for each entry of the input vector. The input-output relation of the output layer is obtained following the same logic, substituting the layer parameters, as well as the ReLU activation by the sigmoid function.

With $\boldsymbol{\theta}_{\text{DFN}}$ in hands, it is possible to calculate the number of trainable parameters in the DFN algorithm. Considering the dimensions of the weights matrices and the bias vectors values, the number of trainable parameters is given by:

$$\Theta_{\text{DFN}}(N_n, L, K) = N_n^2 + 2N_nL + N_nK + 2N_n + K \tag{17}$$

Since the outputs of the DFN are in the range $[0, 1]$, a hard decision module with threshold parameter $\tau \geq 0$ is positioned at the output of the algorithm to calculate the activity descriptor in its original domain. Hence, the hard decision output is:

$$\hat{\alpha}_k = \begin{cases} 0, & \tilde{\alpha}_k < \tau \\ 1, & \tilde{\alpha}_k \geq \tau \end{cases} \tag{18}$$

During the activity detection, two different kinds of error may occur. First, a *false alarm* (FA) occur when an inactive device is detected as active. Second, a *miss detection* (MD) occur when an active device is detected as inactive. The probabilities of FA and MD are calculated in terms of the *hard decision threshold* by:

$$P_{\text{FA}}(\tau) = \Pr(\tilde{\alpha}_k \geq \tau \,|\, \alpha_k = 0) \tag{19}$$

$$P_{\text{MD}}(\tau) = \Pr(\tilde{\alpha}_k < \tau \,|\, \alpha_k = 1) \tag{20}$$

Given the imbalance on the number of active devices during each random access slot, the frequency of each type of error changes. The error probability as a function of τ is written in terms of the FA and MD probabilities as

$$P_{\text{E}}(\tau) = (1 - p_a) P_{\text{FA}}(\tau) + p_a P_{\text{MD}}(\tau) \tag{21}$$

The hard decision module can be optimized in order to meet different design criteria, *e.g.* minimize a specific type of error, or a metric which combines the two types of error with different weights. In our work, we evaluate the algorithms adjusting the hard decision thresholdsuch that

$$\tau^* = \underset{\tau \geq 0}{\arg\min}\, P_{\text{E}}(\tau) \tag{22}$$

B. Convolutional Neural Network

Now, we introduce the CNN architecture depicted in the Fig. 4. Differently from the DFN, the CNN algorithm is implemented with convolutional layers. Particularly, we need to adjust the input dimension aiming to match it with the number of outputs. In order to accomplish this while exploiting the good correlation properties of the ZC preamble sequences, we use a correlator stage at the input of the network which produces the signal

$$\mathbf{y}_{\text{corr}} = \mathbf{A}^H \mathbf{y} \tag{23}$$

This correlator stage can be seen as an additional layer with fixed weights instead of trainable parameters. Then, the inputs of the CNN are the real and imaginary parts of the correlated received signal, organized as two vectors of dimension $K \times 1$. The output is an estimate of the activity descriptor.

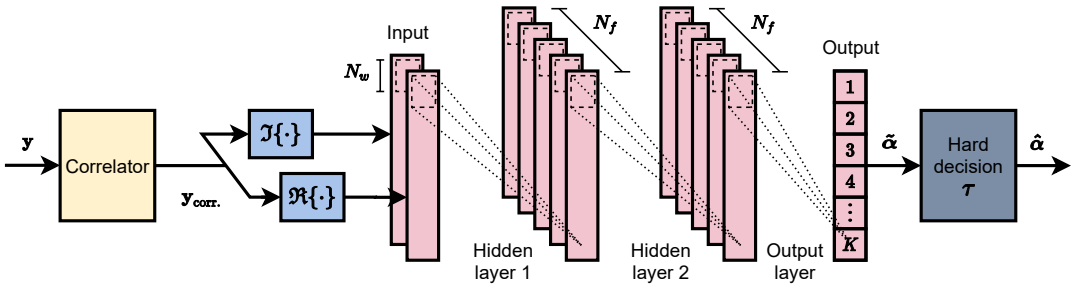


Figure 4. CNN algorithm, which inputs are the real and imaginary parts of the correlated received signal at the BS.

Table II
DESCRIPTION OF THE LAYERS IN THE CNN ALGORITHM.

Layer	Type	Activation	Feature maps	Filter length
Hidden 1	1D Conv	ReLU	N_f	N_w
Hidden 2	1D Conv	ReLU	N_f	N_w
Output	1D Conv	Sigmoid	1	N_w

Conv: convolution.

The CNN architecture has two hidden layers of the type 1D convolution with N_f feature maps each one, filters of length N_w and ReLU as the activation functions. The output layer has one feature map, filters of length N_w and sigmoid as the activation function. The details on the layers of the CNN algorithm are described in Table II.

The parameters of the CNN layers are the filters and the bias values with respect to each feature map. Let the superscript ℓ be the index of the feature map. The parameters of the first hidden layer are $\mathbf{W}_1^\ell \in \mathbb{R}^{N_w \times 2}$ and $b_1^\ell \in \mathbb{R}, \ell = 1, \ldots, N_f$. Next, the parameters of the second hidden layer are $\mathbf{W}_2^\ell \in \mathbb{R}^{N_w \times N_f}$ and $b_2^\ell \in \mathbb{R}, \ell = 1, \ldots, N_f$. Finally, the parameters of the output layer are $\mathbf{W}_3^1 \in \mathbb{R}^{N_w \times N_f}$ and $b_3^1 \in \mathbb{R}$. With these definitions, we assume the set with the CNN parameters as:

$$\boldsymbol{\theta}_{\text{CNN}} = \left\{ \{\mathbf{W}_1^\ell, b_1^\ell\}_{\ell=1}^{N_f}, \{\mathbf{W}_2^\ell, b_2^\ell\}_{\ell=1}^{N_f}, \{\mathbf{W}_3^1, b_3^1\} \right\} \tag{24}$$

Given the respective dimensions of the CNN parameters, we calculate the number of trainable parameters in the algorithm by the equation:

$$\Theta_{\text{CNN}}(N_f, N_w) = N_f^2 N_w + 3N_f N_w + 2N_f + 1 \tag{25}$$

It is important to stress that, differently from the DFN, the number of trainable parameters in the CNN does not depend explicitly on the number of devices and the preamble length. This is a good advantage of the CNN, as its size do not grow rapidly with the scenario parameters. However, N_f and N_w must follow the changes of the parameters in the scenario in order to achieve efficient results on activity detection.

$$[\mathbf{M} * \mathbf{N}]_t = \sum_{i=t-n}^{t-1} \bar{\mathbf{M}}_{i,:} \mathbf{N}_{t-i,:}^T \tag{26}$$

where $\bar{\mathbf{M}} \in \mathbb{R}^{n \times p}$ is a zero-padded version of $\mathbf{M}$. The input-output relation of the hidden layers is given by the equation:

$$[\mathbf{v}_i^\ell]_t = \text{ReLU}\left([\mathbf{W}_i^\ell * \mathbf{V}_{i-1}]_t + b_i^\ell \right) \tag{27}$$

for $i = 1, 2$ and $\ell = 1, \ldots, N_f$. We consider $\mathbf{v}_i^\ell$ as the output of the ℓ-th feature map in the layer i, while $\mathbf{V}_i$ is a matrix which columns are the output vectors of all the feature maps, or channels, in the layer i,

$$\mathbf{V}_i = \begin{bmatrix} \mathbf{v}_i^1 & \cdots & \mathbf{v}_i^{N_f} \end{bmatrix} \tag{28}$$

The input-output relation of the output layer can be obtained easily by substituting the layer parameters and the ReLU activation by the sigmoid. As the DFN, the CNN has a hard decision module to calculate the estimate of the activity descriptor in the original domain.

IV. NUMERICAL RESULTS

In this section, we cover the numerical results associated to the proposed DL algorithms and the preambles design, as well as the baseline techniques for activity detection.

A. DL Training Procedure

The steps required to train both DFN and CNN DL algorithms for solving activity detection problem in mMTC scenarios are described herein. First, observe that a sample represents a transmission interval, storing the real and imaginary parts of the received signal, as well as the devices activity descriptor. We create the samples generating the activity descriptor and evaluating eq. (2). It is worth mentioning that we generate all the signals by sampling random distributions following the definitions in Section II. A data set is constituted by $5 \cdot 10^5$ samples, from which $4.5 \cdot 10^5$ are used for training and $5 \cdot 10^4$ used for validation. Each data set is generated with a single realization of the preamble matrix. The DL algorithms are trained with the data set samples using the adaptive moment (ADAM) estimation algorithm. The training stops if the *loss function* $\mathcal{L}$ in eq. (12) do not improve after 5 epochs. Table III contains the information on the training setup of the DL algorithms.

Table IV contains the average training time of the proposed DL algorithms for different preamble lengths. The training is performed on a workstation equipped with a Nvidia GeForce 940MX GPU, an Intel(R) Core(TM) i5-7200U CPU @2.5 GHz and 8 GB of RAM.

B. DL Input Parameters Tuning

This subsection describes the adopted procedure to tune the architectural parameters of the DL algorithms. We tune the number of neurons in the hidden layers of the DFN and the number of feature maps in the hidden layers of the CNN in order to get good performance on activity detection, while preventing overfitting. The parameters are tuned by hand in the the ranges $N_n \in \{80, 160, 320, 400, 800, 1600\}$ and $N_f \in \{20, 40, 80\}$. The tuning procedure consists on

Table III
TRAINING SETUP USED FOR THE DEEP LEARNING ALGORITHMS VIA ADAPTIVE MOMENT ESTIMATION.

Training algorithm	ADAM estimation
Training samples	$4.5 \cdot 10^5$ (90%)
Validation samples	$0.5 \cdot 10^5$ (10%)
Batch size	128
Maximum epochs	10^5
Epochs for early stopping	5

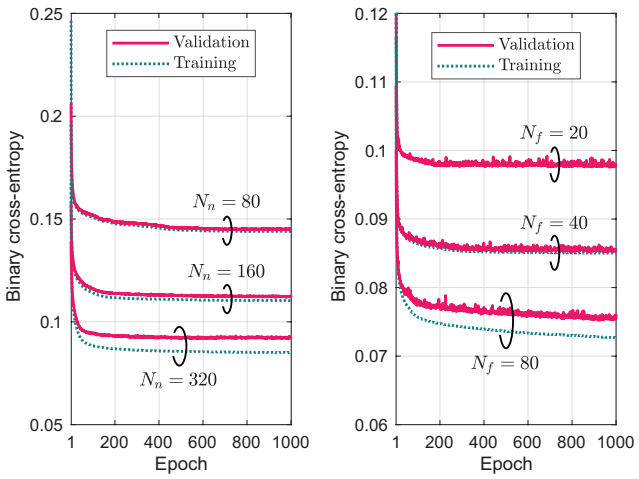
Table IV
TRAINING TIME (SECONDS) OF THE DEEP LEARNING ALGORITHMS ACCORDING TO THE PREAMBLE LENGTH.

L	**DFN**	**CNN**
7	$2.49 \cdot 10^2$	$1.25 \cdot 10^3$
13	$3.88 \cdot 10^2$	$1.72 \cdot 10^3$
17	$3.97 \cdot 10^2$	$1.78 \cdot 10^3$
21	$3.82 \cdot 10^2$	$1.38 \cdot 10^3$

generating and assessing the learning curves of the algorithms, which contain the loss values along the epochs, calculated by the binary cross-entropy function in eq. (12) on the training and validation data sets. Ideally, the loss values calculated for the validation set must follow the values calculated for the training one, indicating improvement on the generalization capability of the algorithm. If the validation loss worsens, or the gap between it and the training loss increases, there is no more gain on increasing the parameter value.

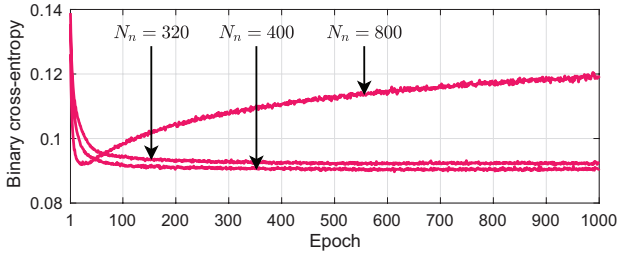
Fig. 5 depicts the learning curves for the DL algorithms varying their architectural parameters[1]. Analyzing the curves, we observe that the minimum training loss always reduces increasing the parameters value, as expected. Fig. 5(a) illustrates the learning curves of the DFN for $N_n \in \{80, 160, 320\}$, while the Fig. 5(c) depicts the validation loss for $N_n \in \{320, 400, 800\}$. Combining these results, one can infer that the DFN attains overfitting with $N_n > 400$. Furthermore, we observe that the validation loss improves marginally from $N_n = 320$ to $N_n = 400$. So, we choose $N_n = 320$ for the DFN algorithm. On the other hand, the validation loss of the CNN, depicted in Fig. 5(b), always decreases with the epochs. Despite that, the gap between the validation and training losses becomes bigger for $N_f > 40$. Considering these facts, we choose $N_f = 40$ for the CNN.

[1]We omit a few curves in order to preserve the readability of the result.



(a) DFN varying the number of neurons in the hidden layers

(b) CNN varying the number of feature maps



(c) Demonstration of overfitting in the DFN

Figure 5. Learning curves with the training and validation losses (binary cross-entropy) versus the epochs of the deep learning algorithms varying their architectural parameters.

C. Performance: Error rate, FA and MD rates

Numerical results for the performance of the proposed DL-based activity detection algorithms for GF-RA protocols are assessed and compared with two reference methods available in the literature, the Least Absolute Shrinkage and Selection Operator (LASSO) method [20] and the AMP algorithm [21]. The adopted AMP algorithm is implemented with the *complex soft thresholding denoising function* [22], while the LASSO method is computed via the *coordinate descent optimization* (CD) algorithm [23]. Both thresholding factor of AMP and regularization parameter of the LASSO are tuned for each evaluated scenario in order to develop a fair comparison, in which the reference schemes achieve good performance-complexity tradeoffs. Three figure of merit are evaluated: *error rate*, *false alarm rate* and *miss detection rate*. Table V summarizes the setup used to produce the evaluated scenarios and respective numerical results. All the signals are generated by sampling random distributions following the definitions in Section II.

Fig. 6 depicts the error rate versus the *undersampling ratio* $\Delta = L/K$ for both DL algorithms, as well as the reference schemes. We evaluate each DL algorithm using the three types of sequences. The first observation is that the detection error rate improves by increasing Δ, as the longer preamble sequences provide more information to the activity detection algorithms. The DFN algorithm using random sequences achieves better performance than the CNN under the same condition. However, the CNN with the ZC sequences achieves detection error rate values comparable to the DFN. The CNN has a huge improvement in detection error rate with the ZC sequences instead of the random ones. This is due the correlator stage in the CNN combined with the good cross-correlation properties of the ZC sequences. The *error rate* performance for the CNN with random sequences are comparable, while the differences on the error rate of the DFN with each type of sequences are non-negligible. Besides, comparing

Table V
SIMULATION PARAMETERS.

Parameter	Value
System	
Number of devices	$K = 40$
Activation probability	$p_a \in [0.01, 0.3]$
Type of sequences	Normal, Bernoulli, Zadoff-Chu
Preamble length	$L \in \{7, 13, 17, 21\}$
Undersampling ratio	$\Delta = L/K \in [0.175, 0.525]$
Preamble power	$\|\mathbf{a}_k\|_2^2 = L, \forall k$
Signal-to-noise ratio	SNR = 10 dB
Channel	
Channel model	Rayleigh fading
Noise power	$\sigma_z^2 = 0.1$
DL algorithms	
Number of neurons	$N_n = 320$
Number of feature maps	$N_f = 40$
Filter length	$N_w = 8$
LASSO [20] via CD	
Maximum number of iterations	$T_{\text{CD}} = 10^3$
Convergence criterion	$\|\hat{\mathbf{x}}^t - \hat{\mathbf{x}}^{t-1}\|_2 \leq 10^{-4}$
AMP [21]	
Denoising function	Complex soft thresholding [22]
Number of iterations	$T_{\text{AMP}} = 10$

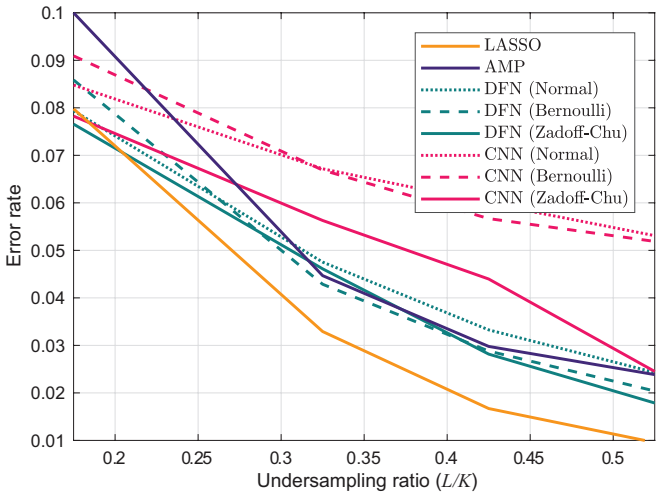


Figure 6. Detection error rate versus the preamble length of the deep learning algorithms using the Normal, Bernoulli and Zadoff-Chu sequences. $p_a = 0.1$.

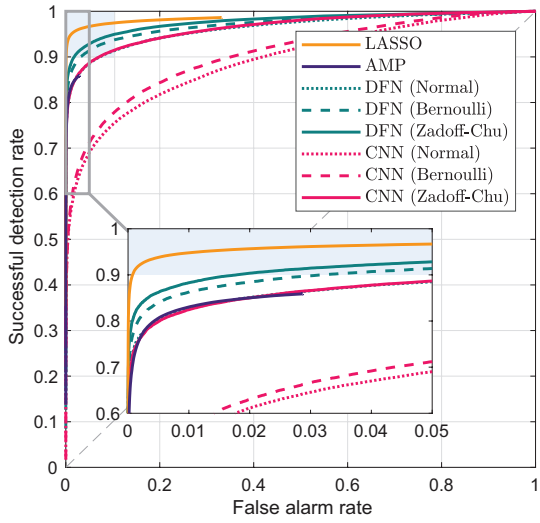
with the reference schemes, the DFN achieves error rate levels comparable to the AMP for $\Delta \geq 0.325$. On the other hand, considering low undersampling ratios, *i.e.,* $\Delta \leq 0.175$ the DFN and the CNN with ZC sequences outperform significantly the AMP scheme. The LASSO method presents the best performance for $\Delta > 0.175$ at the expense of an extremely high computational complexity, as we demonstrate in the subsection IV-E. Moreover, the behavior of the error curves for the evaluated algorithms is similar for $p_a < 0.1$, with decreased error rates as a consequence of the reduced number of active devices, and therefore the level of interference.

Fig. 7 depicts the receiver operating characteristic (ROC)[2] curves of the DL algorithms using the Normal, Bernoulli and ZC sequences types for two values of undersampling ratio. For $\Delta = 0.525$, the CNN using the ZC sequences has a huge improvement on the miss detection rate, if compared with the random sequences. Again, this is due to the correlator stage at the CNN input and the good cross-correlation properties of the ZC sequences. Additionally, the ROC curve of the CNN using the ZC sequences is tight to the AMP algorithm. The LASSO method attains the best ROC curve among the techniques. On the other hand, for $\Delta = 0.175$ the differences between the algorithms and types of sequences vanishes, as all the ROC curves of the DL algorithms are comparable to the LASSO. At the same time, the AMP detection rate performance degrades significantly with Δ increasing.

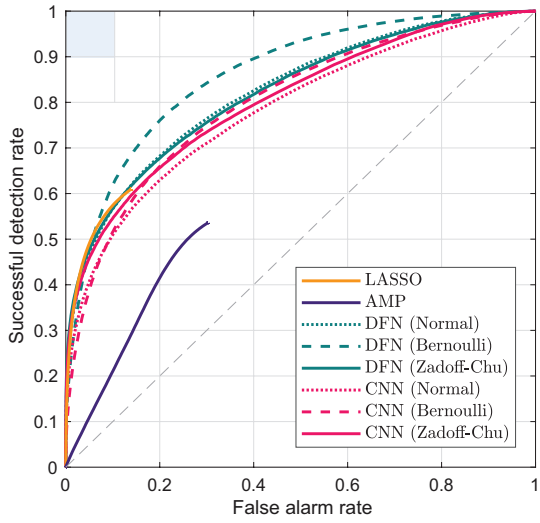
Fig. 8 depicts the detection error tradeoff (DET)[3] curves of the proposed DL algorithms and the reference schemes for both values of undersampling ratio. The proposed DL algorithms cover a wide range of false alarm and miss detection rates, differently from the LASSO and AMP schemes.

[2]The ROC curves are described by the successful detection rate, or true positive rate, versus the false alarm rate, or false positive rate, allowing a comparison of the detection algorithms according to the two types of error.

[3]The DET curves are described by the miss detection rate *vs.* false alarm rate with the axis warped by the inverse of the Normal cumulative distribution function. The DET is more suitable than the ROC to analyze the tradeoff between the two types of error.



(a) $\Delta = 0.525$



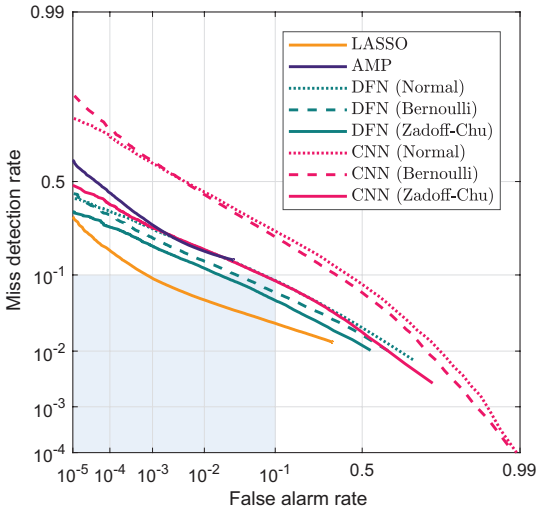
(b) $\Delta = 0.175$

Figure 7. Receiver operating characteristic curves of the DL algorithms using the Normal, Bernoulli and Zadoff-Chu sequences, as well as the LASSO and AMP reference schemes. $p_a = 0.1$.

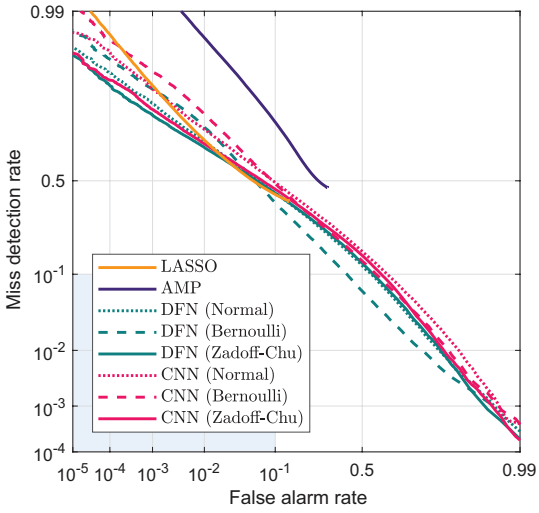
Analyzing the DET curves, we see that the activity detection algorithms achieve lower false alarm rate values than miss detection ones. Such fact is a consequence of the sporadic activity of machine-type devices, which results in more inactive devices than active ones in the random access slots.

D. Robustness of the DFN and CNN Algorithms

Fig. 9 depicts the error rate versus the activation probability of the DL and the reference methods. We evaluate each algorithm using the three types of sequences and two values of preamble length. Such figure of merit demonstrates the robustness of the DL-based activity detection algorithms against the system scenario variations, since the networks are trained for a single-fixed activation probability, $p_a = 0.1$. As expected, the error rate increases with the activation probability, as high numbers of active devices incur in increased interference



(a) $\Delta = 0.525$



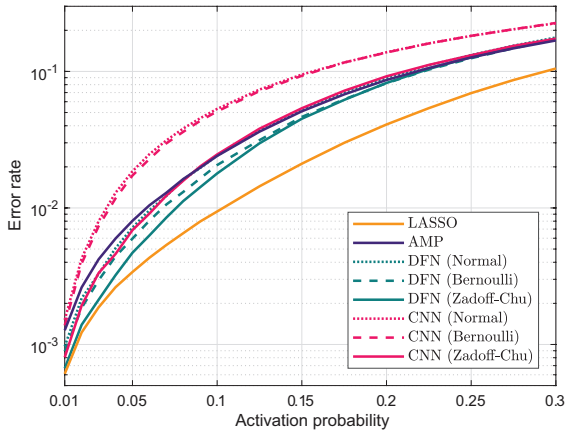
(b) $\Delta = 0.175$

Figure 8. Detection error tradeoff curves for both proposed DL-based algorithms under three distinct preamble sequence sets, as well as the LASSO and AMP reference schemes.

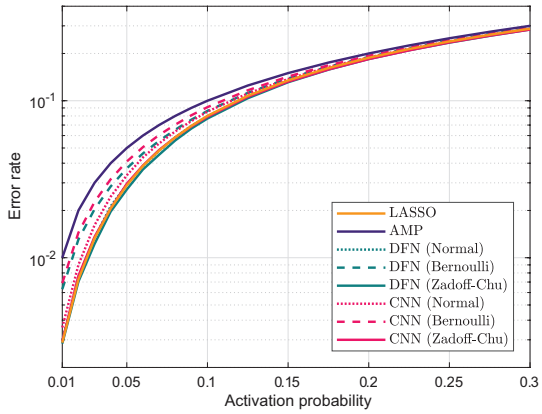
power at the received signal. Moreover, regarding the type of sequences, for $\Delta = 0.525$, again the Zadoff-Chu set achieves the best performance. Both DFN and CNN algorithms present similar results using such sequence sets. Next, the Bernoulli sequences outperform marginally the Normal sets. Specially, the CNN results in significant improvements using the ZC sets, achieving detection error rates close to the DFN. Lastly, when the undersampling ratio is reduced substantially, *e.g.* $\Delta = 0.175$, the performance of the DFN and the CNN algorithms operating under each one of the three types of sequences is tight for $p_a \geq 0.1$.

E. Complexity of the DL-based and Reference Algorithms

Fig. 10 depicts the run time in seconds versus the number of users of the evaluated algorithms for two values of undersampling ratio. In this result, we set $L \in \{71, 211\}$, and



(a) $\Delta = 0.525$



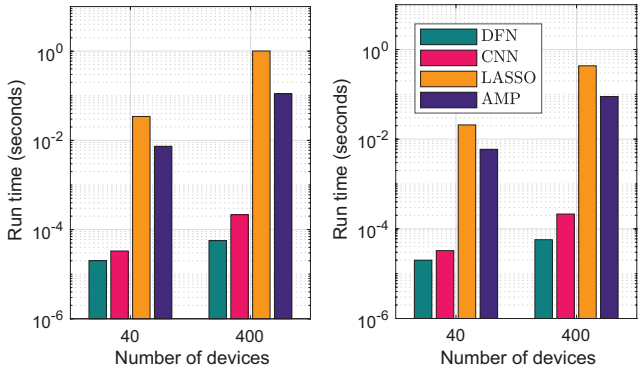
(b) $\Delta = 0.175$

Figure 9. Error rate versus the activation probability of the proposed DL-based activity detection algorithms using the Normal, Bernoulli and Zadoff-Chu sequence sets, as well as the LASSO and AMP reference schemes.

$K \in \{40; 400\}$. The DL-based algorithms present extremely-low run time values on the order of 10 μs to 100 μs. The DFN algorithm has slightly shorter run time than the CNN. Compared with the reference schemes, the proposed algorithms result in run time values at least two orders of magnitude less. As expected, the complexity of all the algorithms increases with the number of devices owing to the growth in the number of inputs. Despite that, the run time changes marginally with the undersampling ratio, except for the LASSO method. Such fact is due to the accelerated convergence of the CD algorithm caused by the amount of information provided by longer preamble sequences.

V. Conclusions

In this work, we propose two DL sparse support recovery algorithms to enable activity detection of machine-type devices in GF-RA protocols. We propose a DFN algorithm for activity detection based on densely connected layers and a CNN algorithm built with 1D convolution layers. At the same time, we analyze the impact of the type of sequences used for preamble design on the performance of the activity detection scheme. We evaluate the performance using random preambles generated by a complex Normal and a Bernoulli distributions,



(a) $\Delta \approx 0.17$ (b) $\Delta \approx 0.52$

Figure 10. Run time *vs.* the number of devices for both DL-based algorithms, and for reference schemes (LASSO, AMP), for two undersampling ratio values.

as well as by deploying deterministic ZC sequences. The numerical results demonstrate that the DFN reaches the best detection error rate values. However, the CNN with the ZC sequences in place of random ones achieves detection error rates comparable to the DFN, due to the good cross-correlation properties of these sequences. On the matter of the computational complexity, both proposed DL-based algorithms present extremely low execution time, at the expense of the training burden to determine the suitable parameter values of the networks. Moreover, the proposed algorithms attain better performance-complexity tradeoff than state-of-art techniques, such as the LASSO method and the AMP algorithm.

REFERENCES

[1] J. Ding, D. Qu, and J. Choi, "Analysis of non-orthogonal sequences for grant-free RA with massive MIMO," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 150–160, Oct. 2020.

[2] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, Mar. 2021.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[4] L. Liu and W. Yu, "Massive connectivity with massive MIMO—Part I: Device activity detection and channel estimation," *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2933–2946, Mar. 2018.

[5] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Transactions on Communications*, vol. 66, no. 12, pp. 6164–6175, Aug. 2018.

[6] A. Fengler, S. Haghighatshoar, P. Jung, and G. Caire, "Non-bayesian activity detection, large-scale fading coefficient estimation, and unsourced random access with a massive MIMO receiver," *IEEE Transactions on Information Theory*, pp. 1–1, Mar. 2021.

[7] M. Borgerding, P. Schniter, and S. Rangan, "AMP-inspired deep networks for sparse linear inverse problems," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4293–4308, May 2017.

[8] Y. Qiang, X. Shao, and X. Chen, "A model-driven deep learning algorithm for joint activity detection and channel estimation," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2508–2512, Jul. 2020.

[9] Y. Shi, S. Xia, Y. Zhou, and Y. Shi, "Sparse signal processing for massive device connectivity via deep learning," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, 7-11 Jun. 2020, pp. 1–6.

[10] W. Kim, Y. Ahn, and B. Shim, "Deep neural network-based active user detection for grant-free NOMA systems," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2143–2155, Jan. 2020.

[11] W. Chen, B. Zhang, S. Jin, B. Ai, and Z. Zhong, "Solving sparse linear inverse problems in communication systems: A deep learning approach with adaptive depth," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 4–17, Jan. 2021.

[12] S. Li, W. Zhang, Y. Cui, H. V. Cheng, and W. Yu, "Joint design of measurement matrix and sparse support recovery method via deep auto-encoder," *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1778–1782, Oct. 2019.

[13] W. Zhang, S. Li, and Y. Cui, "Jointly sparse support recovery via deep auto-encoder with applications in MIMO-based grant-free random access for mMTC," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 26-29 May. 2020, pp. 1–5.

[14] A. Mousavi and R. G. Baraniuk, "Learning to invert: Signal recovery via deep convolutional networks," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 19 Jun. 2017, pp. 2272–2276.

[15] D. Chu, "Polyphase codes with good periodic correlation properties (corresp.)," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.

[16] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, Apr. 2010.

[17] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, Cambs., UK: Cambridge University Press, 2012.

[18] B. Popovic, "Generalized chirp-like polyphase sequences with optimum correlation properties," *IEEE Transactions on Information Theory*, vol. 38, no. 4, pp. 1406–1409, Jul. 1992.

[19] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: The MIT Press, 2012.

[20] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, Jan. 1996.

[21] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, Sep. 2009.

[22] A. Maleki, L. Anitori, Z. Yang, and R. G. Baraniuk, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4290–4308, Mar. 2013.

[23] S. J. Wright, "Coordinate descent algorithms," *Mathematical Programming*, vol. 151, no. 1, pp. 3–34, Mar. 2015.

APPENDIX C – Performance of Activity Detection and Channel Estimation Methods for Machine-Type Communications with MIMO Transceiver

Performance of Activity Detection and Channel Estimation Methods for Machine-Type Communications with MIMO Transceiver

João Henrique Inacio de Souza, Taufik Abrão

***Abstract*—The grant-free random access protocols are promising strategies to provide efficiently, connectivity to the large population of nodes in the massive machine-type communications (mMTC) systems. Besides, the transceiver with multiple antennas is one of the technologies that pushed the 5G, providing benefits like high array gain and inter-channel orthogonality. Taking this into account, in this work we introduce four techniques for activity detection and channel estimation for mMTC systems with MIMO transceiver: the minimum mean-squared error (MMSE) estimator; the group least absolute shrinkage and selection operator (LASSO); the covariance-based estimation formulated as a non-negative least-squares (NNLS) optimization problem; and the vector approximate message passing (VAMP) algorithm. Such techniques are evaluated and compared in terms of the normalized mean-squared error (NMSE), and the detection error rate. We demonstrate the main operation regimes in which each technique can benefit from the multiple antennas at the base-station.**

***Index Terms*—Massive machine-type communications, grant-free random access, multiple-input multiple-output (MIMO), channel estimation**

I. Introduction

The massive machine-type communications (mMTC) service support this typical set of nodes. Numerous devices non-controlled directly by humans that have sporadic activation patterns, transmit small packets with low data rates, have low-complexity hardware and are battery-constrained [1]. Addressing such group of devices with the technologies and strategies of the previous generations is challenging, as they were designed around on a paradigm for human-centric applications.

Taking this into account, new strategies emerged. The grant-free random access schemes are efficient methodologies to deal with the sporadic activation of the mMTC nodes. In such schemes, the active devices access the network and transmit their payload without waiting for any permission of the base-station (BS) [2]. Compared with the grant-based, the grant-free random access schemes require low signaling overhead and message exchanging. However, they are susceptible to collisions and interference, needing sophisticated approaches for collision resolution and interference management.

In general, the grant-free random access techniques are classified in compressed sensing (CS)-based and covariance (CV)-based methods [3]. On the matter of the CS-based techniques, we can cite the least absolute shrinkage and selection operator (LASSO) [4] and the approximate message passing (AMP) algorithm [5]. At the same time, the CV-based methods include the activity detection formulated as a non-negative least-squares (NNLS) optimization [6] problem or using maximum likelihood estimation [7].

In particular, there are techniques that exploit the advantages of the multiple antennas at the BS to detect the set of active devices and estimate their respective channel vectors. In this work, we study a group of techniques of such type available in the recent literature. The methods are the *minimum mean-square error* (MMSE) estimator, the *group LASSO* [8], the NNLS [7] and the *vector AMP* [9].

A. Contributions

In this work, we introduce and evaluate promising techniques available in the literature for activity detection and channel estimation in grant-free random access protocols for mMTC systems with MIMO transceiver. The contributions of our work are as follows:

- Introduction and adaptation of the techniques listed in Table I for the activity detection and channel estimation tasks in mMTC, pointing out their respective advantages and weaknesses;
- Systematic evaluation of the introduced methods and a comprehensive comparison between them in order to identify their performance and main operation regimes.

B. Notation

Boldface small $\mathbf{a}$ and capital $\mathbf{A}$ letters represent respectively vectors and matrices. Capital calligraphic letters $\mathcal{A}$ represent finite sets. $\mathbf{I}_n$ denotes the identity matrix of size n. $\{\cdot\}^T$ and $\{\cdot\}^H$ denote, respectively, the transpose and the conjugate transpose operators. $\text{diag}(\cdot)$ denote the diagonal matrix operator. $\mathcal{CN}(\mu, \sigma^2)$ is a circularly symmetric complex Gaussian distribution with mean μ and variance σ^2. $\text{Bernoulli}(p)$ is a Bernoulli distribution with mean p. $\mathbb{E}[\cdot]$ denotes the expectation operator.

Table I
SUMMARY OF THE INTRODUCED ACTIVITY DETECTION AND CHANNEL ESTIMATION TECHNIQUES FOR MMTC SYSTEMS.

Technique	Activity detection	Channel estimation
MMSE		✓
Group LASSO [8]	✓	✓
NNLS [7]	✓	
VAMP [9]	✓	✓

The rest of the paper is organized as follows. In Section II we introduce the model for the mMTC cellular system. After that, Section III describes the evaluated activity detection and channel estimation techniques for mMTC systems context. In Section IV we analyze the numerical experiments for the introduced methods using four different figures of merit. Lastly, Section V provides the final remarks.

II. System Model

We consider the broadband mMTC cellular system in the following. K machine-type users are served by a BS equipped with a MIMO transceiver with M antennas. They activate sporadically with probability $p_a \ll 1$. During the uplink, in the random access slot, each active user transmits an unique pilot sequence followed by the payload. Since the random access slot is less than the coherence time, the channel response is considered flat. The pilot signals transmitted during the pilot phase are used at the BS for activity detection and channel estimation aiming at enabling the detection of the payload.

The pilot sequences have the length of $L < K$ symbols, owing to the limited coherence time and the typically high number of users in the typical mMTC scenario. We define the pilot sequences as the vectors $\mathbf{a}_k \sim \mathcal{CN}\left(\mathbf{0}, \frac{1}{L}\mathbf{I}_L\right), \forall k$. Besides, the pilots are normalized such that $\|\mathbf{a}_k\|_2^2 = 1, \forall k$.

The channel follows the Rayleigh fading model. Let $\mathbf{h}_k \sim \mathcal{CN}\left(\mathbf{0}, \beta_k \mathbf{I}_M\right), \forall k$ be the channel vectors of the users, considering β_k as the large-scale fading coefficient. Additionally, each antenna at the BS is subjected to an additive white Gaussian noise (AWGN) of power σ^2.

Let γ_k be the *activity descriptor* of user k, which is equal to 1 if the user is active during the random access slot, or 0 otherwise. Considering the sporadic activation of the users, it is reasonable to assume that $\gamma_k \sim \text{Bernoulli}(p_a), \forall k$. Besides, let $\mathcal{K}_a \in \mathbb{Z}_+$ be the set with the indices of the active users. The received pilots signal at the BS during the random access slot is equal to

$$\mathbf{Y} = \sqrt{\rho}\sum_{k=1}^{K} \gamma_k \mathbf{a}_k \mathbf{h}_k^T + \mathbf{Z} \tag{1}$$

where $\rho = L\rho_{\text{UL}}$, considering ρ_{UL} the transmitted power per symbol, and $\mathbf{Z} = \begin{bmatrix}\mathbf{z}_1 & \dots & \mathbf{z}_M\end{bmatrix}$ the AWGN such that $\mathbf{z}_m \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}_L\right), \forall m$. The pilot phase is outlined in Fig 1. We can rewrite the received pilots signal in the convenient matrix form,

$$\mathbf{Y} = \sqrt{\rho}\mathbf{A}\mathbf{\Gamma}\mathbf{H}^T + \mathbf{Z} = \sqrt{\rho}\mathbf{A}\mathbf{X} + \mathbf{Z} \tag{2}$$

where

$$\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_K\end{bmatrix} \tag{3}$$

$$\mathbf{\Gamma} = \text{diag}\left(\begin{bmatrix}\gamma_1 & \cdots & \gamma_K\end{bmatrix}^T\right) \tag{4}$$

$$\mathbf{H} = \begin{bmatrix}\mathbf{h}_1 & \cdots & \mathbf{h}_K\end{bmatrix} \tag{5}$$

$$\mathbf{X} = \begin{bmatrix}\mathbf{x}_1 & \cdots & \mathbf{x}_K\end{bmatrix}^T = \begin{bmatrix}\gamma_1\mathbf{h}_1 & \cdots & \gamma_K\mathbf{h}_K\end{bmatrix}^T \tag{6}$$

Considering the received pilots signal in eq. (2), we compute the signal-to-noise ratio (SNR) as

$$\text{SNR}_k = \frac{\mathbb{E}\left[\|\sqrt{\rho}\mathbf{a}_k\mathbf{h}^T\|_{\text{F}}^2\right]}{\mathbb{E}\left[\|\mathbf{Z}\|_{\text{F}}^2\right]} = \frac{\rho M\beta_k}{LM\sigma^2} = \frac{\rho_{\text{UL}}\beta_k}{\sigma^2} \tag{7}$$

considering that the user k is active, *i.e.* $k \in \mathcal{K}_a$.

In the sequence, we formulate the activity detection and channel estimation problems, performed at the BS from the signal received during the pilot phase.

Activity detection: The activity detection problem consists on estimate the indices of the non-zero rows of the matrix $\mathbf{X}$, namely the row support of $\mathbf{X}$, from the signal $\mathbf{Y}$. Therefore, considering the estimate $\widehat{\mathbf{X}}$ of the signal $\mathbf{X}$, the set $\widehat{\mathcal{K}}_a$ of the estimated active users is equal to

$$\widehat{\mathcal{K}}_a = \text{row supp}(\widehat{\mathbf{X}}) = \{k \in \mathbb{Z}_+ \,|\, \widehat{\mathbf{x}}_k \neq \mathbf{0}\} \tag{8}$$

It is worth mentioning that the sporadic activity of the users, *i.e.* $p_a \ll 1$, implies that the matrix $\mathbf{X}$ has much more null rows than non-zero ones.

Channel estimation: The channel estimation problem consists in determining the vectors $\mathbf{x}_k \, \forall k \in \mathcal{K}_a$ from the received signal $\mathbf{Y}$. Hence, the estimated channel vector of the active users is defined as

$$\widehat{\mathbf{h}}_k = \widehat{\mathbf{x}}_k \, \forall k \in \widehat{\mathcal{K}}_a \tag{9}$$

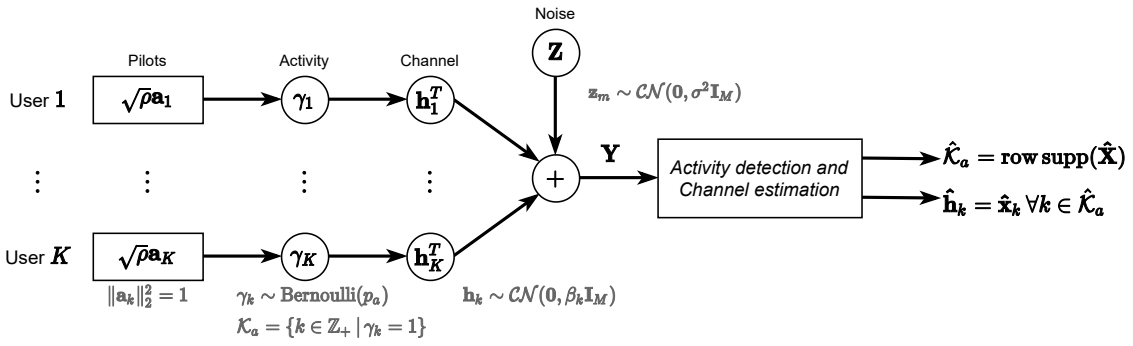


Figure 1. Diagram of the pilot phase, in which the BS perform activity detection and channel estimation. The output of the activity detection and channel estimation module is a set with the indices of the active users and their respective estimated channel vectors.

III. Techniques for Activity Detection and Channel Estimation

In this section, we introduce the techniques for activity detection and channel estimation, especially those methods suitable to operate in mMTC scenarios. We stress that part of the discussed methods are employed only for activity detection, while another part is only used for channel estimation and the rest is employed for both tasks, as stated previously in Table I.

A. Minimum Mean-Squared Error Estimator

The MMSE method is a classical estimator and used for channel estimation in grant-free protocols after the identification of the set of the active users. Hence, given the set of known active users $\widehat{\mathcal{K}}_a$, the MMSE for channel estimation is the matrix that solve the following optimization problem:

$$\mathbf{G}^* = \underset{\mathbf{G}\in\mathbb{C}^{K\times L}}{\arg\min}\, \mathbb{E}\left[\|\mathbf{X}-\mathbf{G}\mathbf{Y}\|_{\mathrm{F}}^2 \,\middle|\, \mathbf{\Gamma}=\widehat{\mathbf{\Gamma}}\right] \tag{10}$$

The eq. (10) has a closed-form solution, computed for each active user $k \in \widehat{\mathcal{K}}_a$ as

$$\mathbf{g}_k^* = \left[\frac{1}{\sqrt{\rho}}\left(\mathbf{A}^H\widehat{\mathbf{\Gamma}}\mathbf{A}+\frac{\sigma^2}{\rho}\widehat{\mathbf{\Gamma}}\mathbf{B}^{-1}\right)^{-1}\widehat{\mathbf{\Gamma}}\mathbf{A}^H\right]_{k,:} \tag{11}$$

Hence, the estimated channel vector of the active users obtained by the MMSE estimator is defined by

$$\widehat{\mathbf{h}}_k^{\mathrm{MMSE}} = (\mathbf{g}_k^*\mathbf{Y})^T, \forall k \in \widehat{\mathcal{K}}_a \tag{12}$$

We stress that the computation of $\widehat{\mathbf{h}}_k^{\mathrm{MMSE}}$ is consistent only for the users that are actually active. In the case of false positive results of the activity detection scheme, a criterion involving the estimated channel vector can be adopted to eliminate these events.

B. Group LASSO

The problem of estimating $\mathbf{X}$ from $\mathbf{Y}$, knowing that $\mathbf{X}$ has a *row-sparse* structure, can be cast as the optimization problem

$$\widehat{\mathbf{X}}^{\mathrm{LASSO}} = \underset{\mathbf{X}\in\mathbb{C}^{K\times M}}{\arg\min}\, \frac{1}{2}\|\mathbf{Y}-\mathbf{A}\mathbf{X}\|_{\mathrm{F}}^2 + \lambda\sum_{k=1}^{K}\|\mathbf{x}_k\|_2 \tag{13}$$

which has the same form of the *group LASSO* [8] problem. The first part of eq. (13) is the MSE between the received signal $\mathbf{Y}$ and its reconstruction from the estimated $\mathbf{X}$. At the same time, the second part induces the sparsity of the rows of $\mathbf{X}$, with the factor λ controlling the sparsity level.

The solution of the problem in eq. (13) can be computed by the block-coordinate descent (BCD) method introduced in [8], and used in the context of grant-free random access in [10]. The BCD method to solve the group LASSO is summarized in Algorithm 1, where N_{it} is the number of iterations and the soft threshold function is defined as

$$\mathrm{soft}(\mathbf{u},\eta) = \begin{cases} \frac{\mathbf{u}}{\|\mathbf{u}\|_2}(\|\mathbf{u}\|_2-\eta), & \text{if } \|\mathbf{u}\|_2 > \eta \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

After the execution of the BCD algorithm, the set of active users is determined by subjecting the rows of the $\widehat{\mathbf{X}}^{\mathrm{LASSO}}$ to a cut-off thresholding approach. The set of detected users is calculated by

$$\widehat{\mathcal{K}}_a^{\mathrm{LASSO}} = \{k \in \mathbb{Z}_+ \,|\, \|\widehat{\mathbf{x}}_k^{\mathrm{LASSO}}\|_2 > \epsilon\} \tag{15}$$

where ϵ is the threshold value. Finally, the estimated channel vector is obtained by

$$\widehat{\mathbf{h}}_k^{\mathrm{LASSO}} = \widehat{\mathbf{x}}_k^{\mathrm{LASSO}}, \quad \forall k \in \widehat{\mathcal{K}}_a^{\mathrm{LASSO}} \tag{16}$$

Algorithm 1: BCD method for the group LASSO [10]

Input: $\mathbf{A}, \mathbf{Y}, \lambda, N_{it}$
Output: $\widehat{\mathbf{X}}^{\mathrm{LASSO}}$

$\widehat{\mathbf{X}}^{(0)} \leftarrow \mathbf{0}$;
for $n = 1, \ldots, N_{it}$ **do**
  for $k = 1, \ldots, K$ **do**
    $\boldsymbol{\chi} \leftarrow \|\mathbf{a}_k\|_2^2\widehat{\mathbf{x}}_k^{(n-1)} - \mathbf{a}_k^H(\mathbf{A}\widehat{\mathbf{X}}^{(n-1)} - \mathbf{Y})$;
    $\widehat{\mathbf{x}}_k^{(n)} \leftarrow \mathrm{soft}\left(\frac{\boldsymbol{\chi}}{\|\mathbf{a}_k\|_2^2}, \frac{\lambda}{\|\mathbf{a}_k\|_2^2}\right)$;
$\widehat{\mathbf{X}}^{\mathrm{LASSO}} \leftarrow \widehat{\mathbf{X}}^{(n)}$;

C. Covariance-based Activity Detection

The covariance-based activity detection methods are based on the covariance matrix of the received signal $\mathbf{Y}$. We introduce a technique in which statistical knowledge of the channel is assumed, performing the activity detection by minimizing the difference between the true covariance matrix and an estimate obtained from the received signal.

In our context, the covariance matrix of the received signal can be defined as

$$\mathbf{\Sigma} = \frac{1}{M}\mathbb{E}\left[\mathbf{Y}\mathbf{Y}^H\right] = \rho\sum_{k=1}^{K}\gamma_k\beta_k\mathbf{a}_k\mathbf{a}_k^H + \sigma^2\mathbf{I}_L, \tag{17}$$

considering the expectation calculated w.r.t. $\mathbf{H}$ and $\mathbf{Z}$. Letting $\boldsymbol{\theta} = \begin{bmatrix}\gamma_1\beta_1 & \cdots & \gamma_K\beta_K\end{bmatrix}^T$, one can conveniently rewrite the covariance matrix as:

$$\mathbf{\Sigma}(\boldsymbol{\theta}) = \rho\mathbf{A}\mathrm{diag}\left(\boldsymbol{\theta}\right)\mathbf{A}^H + \sigma^2\mathbf{I}_L \tag{18}$$

The covariance matrix of the received signal can be estimated by computing

$$\widehat{\mathbf{\Sigma}} = \frac{1}{M}\mathbf{Y}\mathbf{Y}^H \tag{19}$$

The set of active users can be predicted by solving the NNLS optimization problem [7]:

$$\widehat{\boldsymbol{\theta}}^{\mathrm{NNLS}} = \underset{\boldsymbol{\theta}\geq 0}{\arg\min}\, f(\boldsymbol{\theta}) \tag{20a}$$

$$f(\boldsymbol{\theta}) = \|\mathbf{\Sigma}(\boldsymbol{\theta}) - \widehat{\mathbf{\Sigma}}\|_{\mathrm{F}}^2 \tag{20b}$$

The problem in (20) can be solved by gradient coordinate-wise methods, finding the global optimizer for each component of $\boldsymbol{\theta}$ independently. Hence, computing the gradient of the function in (20b), considering a step equal to d_k in the k-th component of $\boldsymbol{\theta}$, we get

$$\frac{\partial f(\boldsymbol{\theta}+d_k\mathbf{e}_k)}{\partial\theta_k} = 2\rho^2 d_k\|\mathbf{a}_k\|_2^4 + 2\rho\mathbf{a}_k^H\left[\mathbf{\Sigma}(\boldsymbol{\theta}) - \widehat{\mathbf{\Sigma}}\right]\mathbf{a}_k \tag{21}$$

Algorithm 2: CD method for the NNLS [7]

Input: $\mathbf{A}, \widehat{\boldsymbol{\Sigma}} = \frac{1}{M}\mathbf{Y}\mathbf{Y}^H, \sigma^2, \beta_k \forall k, N_{it}$
Output: $\widehat{\boldsymbol{\theta}}^{\text{NNLS}}$

1 $\widehat{\boldsymbol{\theta}}^{(0)} \leftarrow \mathbf{0}$;
2 $\boldsymbol{\Sigma} \leftarrow \sigma^2 \mathbf{I}_L$
3 **for** $n = 1, \ldots, N_{it}$ **do**
4 **for** $k = 1, \ldots, K$ **do**
5 $d \leftarrow \frac{\mathbf{a}_k^H \left(\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\right)\mathbf{a}_k}{\rho \|\mathbf{a}_k\|_2^4}$;
6 $d^\star \leftarrow \min\left(\max\left(d, -\widehat{\theta}_k^{(n-1)}\right), \beta_k - \widehat{\theta}_k^{(n-1)}\right)$;
7 $\widehat{\theta}_k^{(n)} \leftarrow \widehat{\theta}_k^{(n-1)} + d^\star$;
8 $\boldsymbol{\Sigma} \leftarrow \boldsymbol{\Sigma} + d^\star \mathbf{a}_k \mathbf{a}_k^H$;
9 $\widehat{\boldsymbol{\theta}}^{\text{NNLS}} \leftarrow \widehat{\boldsymbol{\theta}}^{(n)}$;

where $\mathbf{e}_k$ is the standard unit vector w.r.t. the component k. Hence, the optimal step is easily found as:

$$d_k(\boldsymbol{\theta}) = \frac{\mathbf{a}_k^H \left[\widehat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}(\boldsymbol{\theta})\right] \mathbf{a}_k}{\rho \|\mathbf{a}_k\|_2^4} \tag{22}$$

The Algorithm 2 summarizes the *coordinate descent* (CD) procedure to compute the solution of the NNLS problem in eq. (20). The algorithm operates as follows. At the line 5, the optimal update for the k-th component is calculated following (22). Next, at the line 6, the box constraint of [7] is applied to ensure that each component $\widehat{\theta}_k$ lies between 0 and β_k. Lastly, at the lines 7 and 8 the estimate and the covariance matrix are updated, respectively.

Similarly to the group LASSO, the set of active users for NNLS is estimated using a cut-off thresholding approach after the execution of the CD method. Therefore, the set of active users is equal to

$$\widehat{\mathcal{K}}_a^{\text{NNLS}} = \{k \in \mathbb{Z}_+ \,|\, \widehat{\theta}_k^{\text{NNLS}} > \epsilon\} \tag{23}$$

D. Vector Approximate Message Passing

The VAMP [9] is a multiple-measurement vector version of the AMP algorithm, which can be used for joint activity detection and channel estimation. The VAMP algorithm implements the following expressions, assuming the initial residual matrix $\mathbf{R}^{(0)} = \mathbf{Y}$ and the initial estimate matrix $\mathbf{X}^{(0)} = \mathbf{0}$

$$\widehat{\mathbf{x}}_k^{(n+1)} = \eta_n \left[\left(\mathbf{R}^{(n)}\right)^H \mathbf{a}_k + \widehat{\mathbf{x}}_k^{(n)}\right] \tag{24}$$

$$\mathbf{R}^{(n+1)} = \mathbf{Y} - \mathbf{A}\widehat{\mathbf{X}}^{(n+1)} + \ldots + \frac{1}{L}\mathbf{R}^{(n)} \sum_{k=1}^{K} \eta_n' \left[\left(\mathbf{R}^{(n)}\right)^H \mathbf{a}_k + \widehat{\mathbf{x}}_k^{(n)}\right] \tag{25}$$

where $\eta_n : \mathbb{C}^{M\times 1} \to \mathbb{C}^{M\times 1}$ is the denoising function, while $\eta_n' : \mathbb{C}^{M\times 1} \to \mathbb{C}^{M\times M}$ is the Jacobian matrix of it.

The AMP theory [5] states that the estimated signal after the matched filtering phase is statistically distributed such that

$$\left(\mathbf{R}^{(n)}\right)^H \mathbf{a}_k + \widehat{\mathbf{x}}_k^{(n)} \sim \mathcal{CN}\left(\mathbf{x}_k, \mathbf{T}_n\right) \tag{26}$$

where $\mathbf{T}_n$ is the covariance of the error during the n-th iteration. As our channel model does not account the spatial correlation between the antennas, the covariance matrix $\mathbf{T}_n$ is diagonal, such that:

$$\mathbf{T}_n = \text{diag}\left(\begin{bmatrix}\tau_{n,1} & \ldots & \tau_{n,M}\end{bmatrix}^T\right) \tag{27}$$

Considering the statistical distribution of the estimated signal after matched filtering of eq. (26), the MMSE denosing function is designed by computing

$$\eta_n(\mathbf{u}) = \mathbb{E}\left[\mathbf{x}_k \,|\, \mathbf{u}\right] \tag{28}$$

where $\mathbf{u}$ is the left side of (26). Hence, the expression of the MMSE denoising function can be defined as [7]:

$$\eta_{n,k}(\mathbf{u}) = \phi_{n,k}(\mathbf{u})\beta_k(\beta_k \mathbf{I}_M + \mathbf{T}_n)^{-1}\mathbf{u} \tag{29}$$

where

$$\phi_{n,k}(\mathbf{u}) = \left\{1 + \frac{1-p_a}{p_a} \prod_{m=1}^{M} \left[\frac{\beta_k + \tau_{n,m}^2}{\tau_{n,m}^2} \exp\left(-\frac{\beta_k |u_m|^2}{\tau_{n,m}^2(\beta_k + \tau_{n,m}^2)}\right)\right]\right\}^{-1} \tag{30}$$

Also, the Jacobian matrix of the MMSE denoising function is equal to [7]

$$\eta_{n,k}'(\mathbf{u}) = \phi_{n,k}(\mathbf{u})\boldsymbol{\Psi}_{n,k} + \left[\phi_{n,k}(\mathbf{u}) - \phi_{n,k}^2(\mathbf{u})\right] \overline{\boldsymbol{\Psi}}_{n,k} \mathbf{u}\mathbf{u}^H \tag{31}$$

where

$$\boldsymbol{\Psi}_{n,k} = \text{diag}\left(\begin{bmatrix}\frac{\beta_k}{\beta_k + \tau_{n,1}^2} & \ldots & \frac{\beta_k}{\beta_k + \tau_{n,M}^2}\end{bmatrix}^T\right) \tag{32}$$

$$\overline{\boldsymbol{\Psi}}_{n,k} = \text{diag}\left(\begin{bmatrix}\frac{1}{\tau_{n,1}^2}\left(\frac{\beta_k}{\beta_k + \tau_{n,1}^2}\right)^2 & \ldots & \frac{1}{\tau_{n,M}^2}\left(\frac{\beta_k}{\beta_k + \tau_{n,M}^2}\right)^2\end{bmatrix}^T\right) \tag{33}$$

After the last iteration of the VAMP algorithm, the set of active users is computed by the cut-off thresholding approach, resulting in

$$\widehat{\mathcal{K}}_a^{\text{VAMP}} = \{k \in \mathbb{Z}_+ \,|\, \|\widehat{\mathbf{x}}_k^{\text{VAMP}}\|_2 > \epsilon\} \tag{34}$$

Then, the estimated channel vectors are equal to

$$\widehat{\mathbf{h}}_k^{\text{VAMP}} = \widehat{\mathbf{x}}_k^{\text{VAMP}}, \quad \forall k \in \widehat{\mathcal{K}}_a^{\text{VAMP}} \tag{35}$$

IV. Numerical Results

In this section, we present the numerical results to evaluate the performance of the proposed techniques for activity detection and channel estimation in mMTC scenarios. The four analyzed techniques are the genie-aided MMSE, the group LASSO and the NNLS. The VAMP technique will be evaluated in the updated version of this manuscript.

A. Figures of Merit

Now, we present the figures of merit adopted to compare the techniques for activity detection and channel estimation. The performance of the activity detection techniques is measured by the *detection error probability*, defined as

$$P_e \triangleq \Pr\left(\widehat{\gamma}_k \neq \gamma_k\right) \tag{36}$$

where γ_k is the activity descriptor. Also, the performance of the activity detection techniques w.r.t. each type of error is measured via the *miss detection* and *false alarm* probabilities, defined respectively as

$$P_{\text{MD}} \triangleq \Pr\left(\widehat{\gamma}_k = 0 \,|\, \gamma_k = 1\right) \tag{37}$$

$$P_{\text{FA}} \triangleq \Pr\left(\widehat{\gamma}_k = 1 \,|\, \gamma_k = 0\right) \tag{38}$$

The probabilities of miss detection and false alarm are associated with the detection error probability by the relation:

$$P_e = p_a P_{\text{MD}} + (1 - p_a) P_{\text{FA}} \tag{39}$$

The quality of the channel estimation is measured via the normalized mean-squared error (NMSE), defined as

$$\text{NMSE} \triangleq \frac{\|\mathbf{X} - \widehat{\mathbf{X}}\|_{\text{F}}^2}{\|\mathbf{X}\|_{\text{F}}^2} \tag{40}$$

The NMSE compares the magnitude of the estimation error w.r.t. the original signal.

B. Details on the Activity Detection Techniques

The evaluated activity detection techniques do not provide explicitly an estimate of the activity descriptor $\boldsymbol{\gamma}$ in its original domain. For this reason, as is discussed in Section III, We use an approach based on a cut-off threshold to compute the estimate of the activity descriptor. Let the cut-off thresholding operator be defined as

$$f(u_k) = \begin{cases} 1, & \text{if } u_k > \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{41}$$

where $u_k, \forall k$ and ϵ are non-negative values. For the techniques that perform joint activity detection and channel estimation (group LASSO and VAMP), $u_k = \|\widehat{\mathbf{h}}_k\|_2$. On the other hand, for the NNLS technique, $u_k = \widehat{\theta}_k$.

Considering the cut-off threshold approach, the miss detection and false alarm probabilities can be written in terms of the threshold ϵ as

$$P_{\text{MD}}(\epsilon) = \Pr\left(u_k \leq \epsilon \,|\, \gamma_k = 1\right) \tag{42}$$

$$P_{\text{FA}}(\epsilon) = \Pr\left(u_k > \epsilon \,|\, \gamma_k = 0\right) \tag{43}$$

Hence, the threshold value that minimizes the detection error probability can be computed by solving the optimization problem

$$\epsilon^* = \underset{\epsilon \geq 0}{\arg\min}\, P_e(\epsilon) \tag{44a}$$

$$P_e(\epsilon) = p_a P_{\text{MD}}(\epsilon) + (1 - p_a) P_{\text{FA}}(\epsilon) \tag{44b}$$

C. Simulation Scenario

To evaluate the presented techniques for activity detection and channel estimation in mMTC, let's consider the communication scenario of Fig. 1. All the signals of eq. (2) are generated by sampling random distributions according to the definitions in Section II. The simulation scenario has the follow configuration. The number of users is $K = 100$ and the activation probability is $p_a = 0.1$. Also, the number of antennas at the BS is $M \in \{4, 64\}$, while the pilot length is $L \in [10, 50]$. The power transmitted by the users is $\rho_{\text{UL}} = 23$ dBm, and the power spectral density (PSD) of the noise is -169 dBm/Hz, while the system bandwidth is equal to 1 MHz. The large-scale fading coefficients in dB are defined by $\beta_k = -128.1 - 36.7 \log_{10}(\varrho_k), \forall k$, where ϱ_k is the distance from user k to the BS in kilometers, distributed randomly in the range $(0.05, 1)$ [9]. All the simulation parameters are organized in Table II.

D. Performance of Channel Estimation

Fig. 2 depicts the NMSE of the channel estimate *versus* the undersampling ratio, defined as L/K, of the genie-aided MMSE and the group LASSO. We consider two values for the number of antennas at the BS. For all the techniques, the NMSE decreases as the undersampling ratio increases, since the longer pilot sequences provide more information to perform channel estimation. The evaluated techniques reach a floor of NMSE after a certain undersampling ratio value. Therefore, there is a limit on the performance gain provided by increasing the pilot sequence length. The genie-aided MMSE achieves the lowest NMSE values, outperforming significantly the group LASSO. However, it is worth to mention that, differently from group LASSO, this technique assumes perfect knowledge of the set of active users and exploits information of the large-scale fading coefficients and the noise power.

Now, we discuss the impact of the number of antennas at the BS on the channel estimation. For the genie-aided MMSE, the impact of the number of antennas on the NMSE is marginal. On the other hand, if compared with 4 antennas, the NMSE of the group LASSO with 64 antennas improves significantly

Table II
SIMULATION PARAMETERS.

Parameter	Value
System	
Number of devices	$K = 100$
Activation probability	$p_a = 0.1$
Pilot length	$L \in [10, 50]$
Undersampling ratio	$\Delta = L/K \in [0.1, 0.5]$
Transmitted power	$\rho_{\text{UL}} = 23$ dBm
Signal-to-noise ratio	$\text{SNR} = 10$ dB
Channel	
Channel model	Rayleigh fading
Noise PSD	-169 dBm/Hz
Bandwidth	1 MHz
Distance to the BS	$\varrho_k \in [0.05, 1], \forall k$
Group LASSO	
Regularization parameter	$\lambda = 1.9 \cdot 10^{-6}$
Number of iterations	$N_{it} = 100$
NNLS	
Number of iterations	$N_{it} = 10$

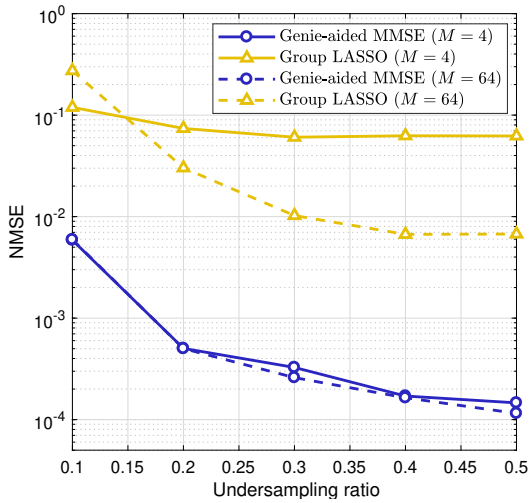


Figure 2. NMSE of the estimated channel vectors of the detected users *vs.* the undersampling ratio (L/K) of the genie-aided MMSE and the group LASSO. Two values for the number of antennas at the BS are considered.

for $\Delta \geq 0.2$. When $\Delta = 0.1$, the NMSE achieved with 64 antennas is greater than the obtained with 4.

E. Performance of Activity Detection

Fig. 3 depicts the detection error rate *versus* the undersampling ratio of the NNLS and the group LASSO. Again, we consider two values for the number of antennas at the BS. We stress that the cut-off threshold to generate this results is calculated by the optimization problem in eq. (44). The performance improves as the undersampling ratio increases. However, the techniques reach a floor of error rate after a certain value of undersampling ratio, which depends on the method and the number of antennas. Except for the case when $\Delta = 0.1$ and $M = 4$, the group LASSO provides remarkably lower error rates if compared with the NNLS. On the matter of the number of antennas at the BS, the NNLS with $M = 64$ achieves better results than with $M = 4$. At the same time, the group LASSO with $M = 64$ improves significantly for $\Delta \geq 0.3$, when compared with the case with $M = 4$. Nevertheless, such behavior reverses for $\Delta < 0.3$.

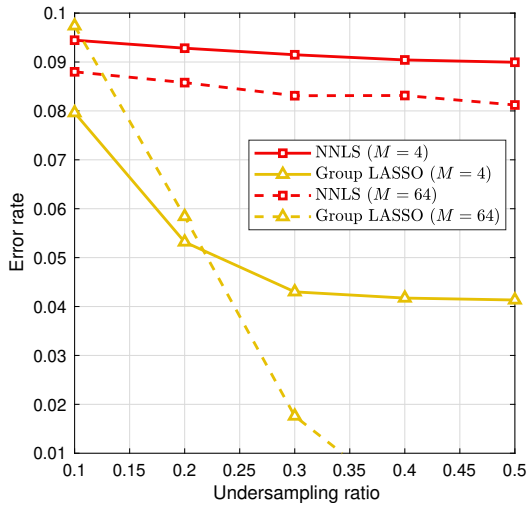
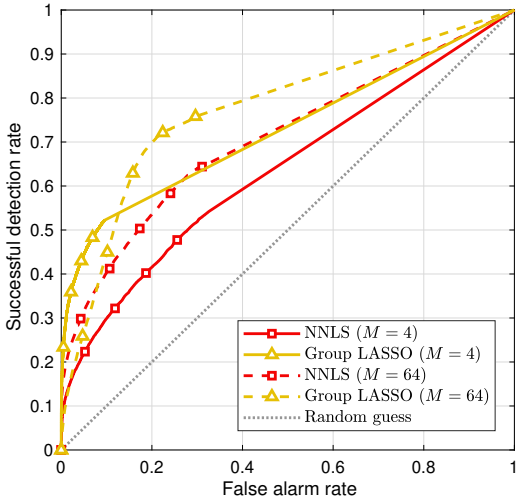
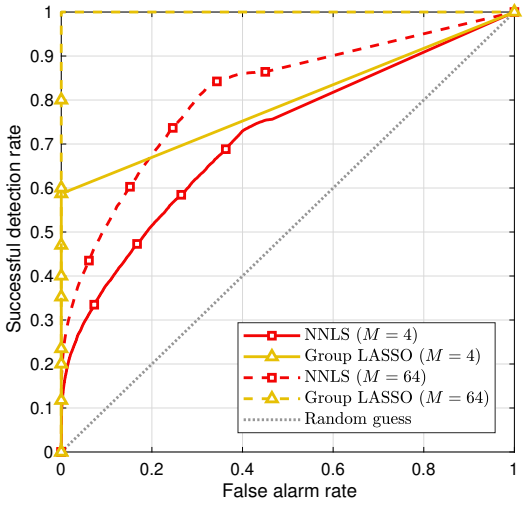


Figure 3. Detection error rate *vs.* the undersampling ratio (L/K) of the NNLS and group LASSO considering two values of antennas at the BS.

In the sequence, Fig. 4 depicts the receiver operating characteristic (ROC) curves of the NNLS and group LASSO for two values of number of antennas at the BS and considering two values of undersampling ratio. The ROC illustrates the trade-off between successful detection rate, the complement of the miss detection rate, and the false alarm rate. Analyzing the results, we note that, under the same conditions, the group LASSO has better ROC curves than the NNLS. Comparing the curves of group LASSO for $\Delta = 0.1$, we see that the technique achieves lower false alarm rates with $M = 4$ and higher successful detection rates $M = 64$. Analyzing the remaining



(a) $\Delta = 0.1$

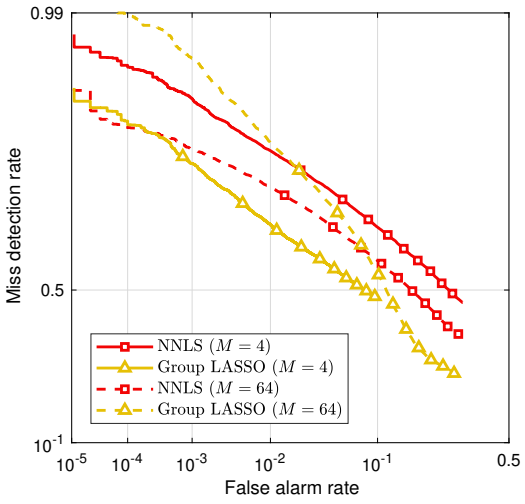


(b) $\Delta = 0.5$

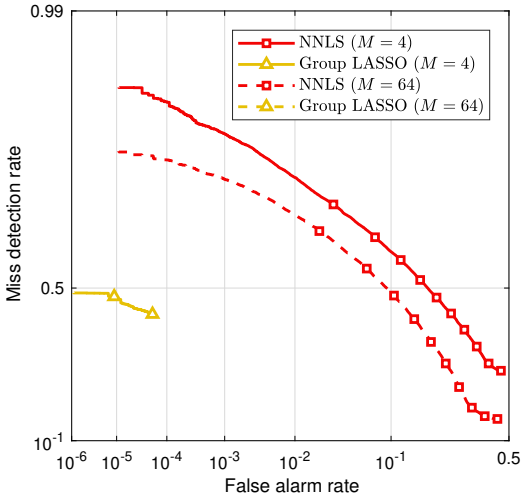
Figure 4. Receiver operating characteristic curves of the NNLS and group LASSO for two values of antennas at the BS and considering two values of undersampling ratio (L/K). The dotted curves represent the performance of choosing the activity descriptors randomly.

curves, we perceive that increasing the number of antennas increases the successful detection rate and decreases the false alarm rate.

Lastly, Fig. 5 depicts the detection error trade-off (DET) curves of the NNLS and the group LASSO with two number of antennas at the BS. Two values for the undersampling ratio are considered. The DET curves provide a better visualization of the trade-off between the two types of error that occur during activity detection than the ROC ones. Analyzing the result, we see that, except for the group LASSO and $\Delta = 0.1$, increasing the number of antennas decreases both the false alarm and the miss detection rates. Specially for $\Delta = 0.1$, the group LASSO with $M = 64$ degrades when compared with $M = 4$. Such fact complements the analysis on the degradation of error rate observed for the group LASSO observed when $\Delta = 0.1$ in Fig. 3.



(a) $\Delta = 0.1$



(b) $\Delta = 0.5$

Figure 5. Detection error trade-off curves of the NNLS and group LASSO for two values of antennas at the BS and considering two values of undersampling ratio (L/K).

V. Conclusions

In this work, we explore activity detection and channel estimation methods for mMTC systems with MIMO transceiver. Four different approaches are introduced and evaluated in terms of NMSE of the estimated channel vectors and the performance on the activity detection, including analysis of the detection error rate, as well as the ROC and DET curves. On the matter of the channel estimation, we demonstrate that the group LASSO technique benefits consistently from increasing the number of antennas, as long as it operates with a sufficiently high pilot length. At the same time, the group LASSO achieves good results on activity detection, outperforming the NNLS technique when the pilot length is high.

References

[1] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: physical and mac-layer solutions," *IEEE Communications Magazine*, vol. 54, pp. 59–65, Sep. 2016.

[2] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the internet of things," *IEEE Signal Processing Magazine*, vol. 35, pp. 88–99, Sep, 2018.

[3] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, pp. 615–637, Mar. 2021.

[4] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, pp. 267–288, Jan. 1996.

[5] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, pp. 18914–18919, Sep. 2009.

[6] C. Wang, O. Y. Bursalioglu, H. Papadopoulos, and G. Caire, "On-the-fly large-scale channel-gain estimation for massive antenna-array base stations," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, 20-24 May 2018.

[7] A. Fengler, S. Haghighatshoar, P. Jung, and G. Caire, "Non-bayesian activity detection, large-scale fading coefficient estimation, and unsourced random access with a massive MIMO receiver," *IEEE Transactions on Information Theory*, pp. 1–1, Mar. 2021.

[8] Z. Qin, K. Scheinberg, and D. Goldfarb, "Efficient block-coordinate descent algorithms for the group lasso," *Mathematical Programming Computation*, vol. 5, pp. 143–169, Jun. 2013.

[9] L. Liu and W. Yu, "Massive connectivity with massive MIMO—Part I: Device activity detection and channel estimation," *IEEE Transactions on Signal Processing*, vol. 66, pp. 2933–2946, Mar. 2018.

[10] S. Li, W. Zhang, and Y. Cui, "Jointly sparse signal recovery via deep auto-encoder and parallel coordinate descent unrolling," in *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2020.