Resource Management for Flexible 5G RAN architecture

Ubaldo Bucci

Department of Information Engineering,
Computer Science and Mathematics

Ph.D. Program in ICT - Systems Engineering telecommunications and HW/SW platforms
XXXIV cycle - SSD ING-INF/05

Università degli Studi dell'Aquila

Advisor: Prof. Dajana Cassioli
.....

Co-Advisor: Prof. Fabio Graziosi
.....

Coordinator: Prof. Vittorio Cortellessa

A thesis submitted for the degree of
Doctor of Philosophy

2022

Contents

Abstract

This Ph.D thesis describes the acquired expertise and the work done during my PhD program started in 2018 and finished in 2022. The research objective is to investigate new resource management schemes in novel *flexible* architectures for the 5G Radio Access Network.

Flexibility can be achieved in several different ways: by flexible architectures, by the softwarization of network functions, by enabling flexible topologies, etc.

In this thesis, we start by reviewing in **chapter 1** the landscape of the newest radio access configurations that have been accepted and deployed to build the new generation of 5G networks. For each of them, we provide details about their architectures. Simultaneously, we provide the benefits stemming from their deployment as radio access network infrastructure.

Then, in **chapter 2**, we show the benefits of the *software programmability* of the network through the network function virtualization and software-defined networking, which opens the way to the new paradigm of network slicing. Network slicing allows the creation of several logical networks on top of a unique physical infrastructure by instantiating a set of virtualized network functions dedicated to specific services and configurations which can be deployed in strategic nodes into the physical network. In particular, the work developed for this thesis focuses on *flexible numerology* and *bandwidth part management*. The flexible approach to management is based on virtualization and software-defined networking paradigms and has been tested in a 5G Radio Access Network exploiting millimetre-waves transmission. This study resulted in a developed allocation algorithm for bandwidth parts which is **SDN-assisted**.

Finally, we analysed the benefits of assuming a flexible network topology through the new cell-free paradigm, which gives major performance improvements when coupled with the massive MIMO technology enabled by high-frequency transmissions, like e.g. the millimeter-waves. Hence,

the **chapter 3** includes all the maths fundamentals which we developed to demonstrate the validity of the Cell-Free Massive MiMo technique in a 5G context, where innovative services with high requirements in terms of QoS, bandwidth and latency, should be supported. We focus on the **ultra reliable low latency communication** and the **enhanced Mobile BroadBand** services and show their performance in a scenario with flexible resource allocation, where the orthogonality of assigned resources is relaxed between the two coexisting services.

The last chapter, i.e., **chapter 4**, presents, as future development work, the use of artificial intelligence, and in particular the reinforcement learning, to optimize the resource allocation of transmission systems. We provided ideas for modelling the considered radio access networks that are presented in this closing chapter with all its theoretical fundamentals. The reinforcement learning-based approaches have been widely used to manage issues such as energy and spectrum allocation efficiency.

Finally, we draw a conclusion from this broad study on flexibility at the level of resource management in 5G radio access networks.

Acknowledgements

During this PhD educational path, I had the chance to develop very high-level expertise related to very advanced topics. Moreover, it allowed me to build networking with very good scientists and good researchers. This path allowed me, also, to travel to very beautiful countries such as Spain and Greece. The main responsible for this magnificent experience is first of all my Ph.D tutor **Dajana Cassioli**, who has always pushed me to the limit of my capabilities to develop a good researcher profile. So, the first thanks go to the "*Prof.*" Then, I would like to say thank you to my colleagues, the Italian and the new Greeks ones. In particular, I would say thank you to **Andrea Marotta** who has been supervising me since 2018. From my internship period to the very last end. Moreover, I would like to dedicate this thesis work to my family that always supported me and in particular to the person that I lost at the beginning of this path. I realised his importance just when I lost him (but this is life, isn't it). This allows me to develop the mindset of appreciating what I have right here right now and enjoying it in each moment of my life. To this list of thanks, I have to also add my friends. In particular, I would like to say thank you to my Italian friends that always supported me both in the bad and good times of this path. I would like to say thanks **Nikos Episkopos** and **Klearchos Palias**. They introduced me to Athens and helped me whenever I needed it. This resulted in a very good friendship. Moreover, I would like to say thank to the newest and Greek couple of friends of mine who make my Greek experience the most beautiful one in my life. There are no words to describe how kind and beautiful people they are. It was surprising to find such very good people. So **thank you very much SIR/MASTER/TEACHER/LITTLE BOSS and most of all MY FRIEND DIONISIS and THE BIG BOSS CHLOE THEODOROPOULOS**. Then thanks to **TANGOPOLIS CREW** for sharing with me a plethora of moments plenty of joy. I hope to see you again in the future. Last but not least, Since I'm sure that I have forgotten somebody,

I would like to thanks whoever hasn't been included and contributed in someway to the accomplishment of this goal. **Now Pameee**

Chapter 1

Flexible 5G Radio Access network architecture

One of the main issues limiting the evolution process toward a new 5G radio access network is the monolithic architecture of the legacy system. Due to the low flexibility degree in the absence of network programmability, the time to in-field deployment of the network innovations or upgrades is too large and often it is not acceptable. To overcome this limitation and introduce the required flexibility for the 5G radio access network, network and software engineers decided to extend the application of virtualised-based solutions, already applied in the core network, even in the radio access segment. Due to this consideration, this chapter has been arranged as follows: in the **sec 1.1: Next Generation of 5G based Radio Access Networks**, we describe all the flexible radio access network architectures that have been developed and deployed eg. Cloud, heterogeneous, virtualized and fog radio access network, in the **sec 1.2: Network Function Virtualization (NFV) and Software Defined Network (SDN) paradigms**, we describe the enabling software mechanisms allowed the development of the aforementioned radio access configuration, in the **sec 1.3: Network Slicing on Radio Access Network** we introduce the new concept of network slicing limited to the radio access segment that can be considered the unexpected breakthrough technology stemming from the combination of flexible radio architecture, network function virtualization and software-defined networking approach. In sec **sec 1.4: New Promising massive MIMO system: The Cell-free massive MIMO** , we introduce the newest radio configuration that has been named Cell-free massive MIMO that is the most promising configuration asset offering huge performance level fetching the expected 5G service radio requirements.

1.1 Next Generation of 5G-based Radio Access Networks

Along with the history of mobile network evolution, every architectural innovation is driven by new requirements of mobile communication services that cannot be satisfied by legacy systems, which have inherited shortages in their outdated architecture. This consideration represents the know situation of the legacy 4G Network (4Gn). This situation pushed the research community to start the evolution process towards the Fifth Generation of Mobile Communication System (5G MCS). The 5G MCS or simply 5G Network (5Gn), aims to deliver a ubiquitous mobile service with enhanced Quality of Service (QoS). This enhanced QoS comes together with the expected and rapid increments of user density, traffic volume, and data rate. This enables new use cases for different vertical industrial applications such as automobiles, public transportation, medical care, energy, public safety, agriculture, entertainment, manufacturing, and so on. This new converting process towards 5Gn has to be supported by a novel technological solution aiming to convert the legacy network infrastructure into something which can sustain those new use cases. The first attractive solution is to redesign the network architecture starting from the rebuilding of the radio access network Radio Access Network (RAN). Different RAN architectures have been designed to fulfil the requirements of 5G-related new uses-cases. In particular, during the evolution process towards 5G RAN, among the presented and studied architectures it is possible to identify the milestones which are **Cloud-RAN (CRAN)**, **Heterogeneous Cloud-RAN (H-CRAN)**, **Virtualized Cloud-RAN (V-CRAN)**, and **Fog Cloud-RAN (F-CRAN)**. Each RAN architecture has different features ranging from flexibility, reduced energy consumption, reduced operations expenditure, and intelligent and optimized resource allocation strategy system giving, in turn, enhanced spectrum efficiency and optimal overall network performance.

1.1.1 CRAN

The **CRAN** has been proposed by IBM under the name Wireless Network Cloud (WNC) as in [1]. Later, it was described with further details in the white paper [2]. According to it and represented by the fig.1.1, the main concept behind CRAN is the separation of all BBUs from their corresponding RRHs and pooling them into a centralized, cloudified, shared and virtualized BBU pool. Every RRH is connected through a fronthaul link to its corresponding BBU pool. Every BBU pool can support

up to tens of RRHs, and be connected through a backhaul link with the core network as in [3]. This kind of radio access network introduced the expected benefits it has

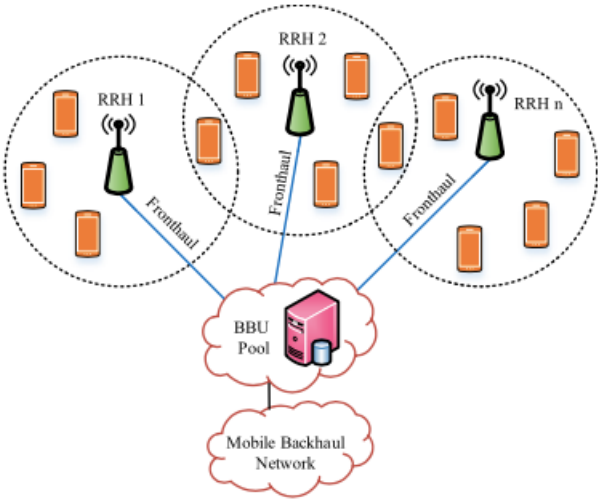


Figure 1.1: CRan from [4]

been designed for. In particular, CRAN, with the centralization of base-band processing, decreased the overall Capital Expenditures (CAPEX) and Operational Expenditures (OPEX) of mobile networks, moreover with its virtualization, the CRAN reduced energy consumption, improved spectral efficiency, and network throughput and facilitated load balancing. In this type of architecture, the functionalities of the radio segments can be divided between Base Band Unit (BBU) and Radio Remote Head (RRH). This give us in turn 2 types of CRAN which are *Fully centralized* and *Partially centralized*. In the first one the functionalities of *Layer-1*, *Layer-2*, *Layer-3* are all located in BBU. In the latter one, the functionalities are split between BBU and RRH. Specifically, *Layer-1*'s functionalities are located at RRH, *Layer-2* and *Layer-3*'s functionalities remain in the BBU. This configuration introduces low transmission bandwidth because the base-band processing has been moved to the RRH while introducing fewer flexibility challenges when the need to upgrade the network comes, as described in [4]. This kind of architecture has been extensively studied from different points of view such as energy consumption, security, performance spectrum sharing, CAPEX and OPEX and mobility management. As shown in [5], [6], [7], the centralization of base-band-processing has reduced the number of BBU and the usage of cooperative radio technologies introduced lower interference between RRHs. These both features give us in turn less energy consumption. From the security point of view in [8] [9] [10] threats together with possible solutions have been identified

and proposed. From the CAPEX and OPEX points of view, in [11] [12] [13] [14], the authors showed that proper deployment strategies for BBU and RRH can reduce the CAPEX and OPEX expenditures and contemporary fulfil the QoS requirements. About performance, the author in [15] [16] [17] shown that CRAN architecture, due to its high flexibility, can introduce great benefits both Dl (increments ranging in 40%-70%) and Ul transmission (the increment is by up 2-3 folds). The evolution towards 5Gn pushed the spectrum to its theoretical limit and on the other side, the deployment of different RAN technologies makes the network dense. This forced the researcher to design new spectrum allocation mechanisms to increase spectrum efficiency. One of the possible ways to increase such efficiency is the deployment of the CRAN as shown in [18] [2]. Here the authors proposed new schemes for spectrum allocation leveraging this architecture to improve spectrum efficiency. Related to the mobility management, CRAN introduced benefits as shown in [19] [20]. In particular in [19] the authors, thanks to CRAN architecture, show how their algorithm can introduce a better QoS, while in [20] CRAN plays a significant role in decreasing of handover delay, moreover, it eliminates the risk of end-user losing its connection.

1.1.2 H-CRAN

Hybrid solutions in the RAN have been designed to limit the side effect of the possibly disruptive converting process towards 5Gn. This means that considering the obtained performance of CRAN, the next step in the evolution process was integrating that architecture with the legacy systems. This led to what we call a hybrid solution and in literature is defined as **H-CRAN**. H-CRAN has been designed to decouple both control and user planes. This enhances the functionalities and performance of CRAN architecture, in which control plane functions are only implemented in the macro BSs. This hybrid solution provides, in turn, both the benefits coming from CRAN architecture and legacy radio access network that is improved spectral efficiency and higher data rates. As depicted in fig.1.2 and presented even in [21] [22], H-CRAN consists of two cellular layouts, the macro BSs (High Power Nodes (HPNs)) cellular layout and the small BSs or RRHs cellular layout. The (HPNs) are mainly deployed to enhance network coverage and control network signalling. Then, we have also the small cells and RRHs which are aimed to guarantee improved network capacity and fulfil the diverse requirements of QoS of various end-users. Also, in this RAN architecture we can identify three main features which are:

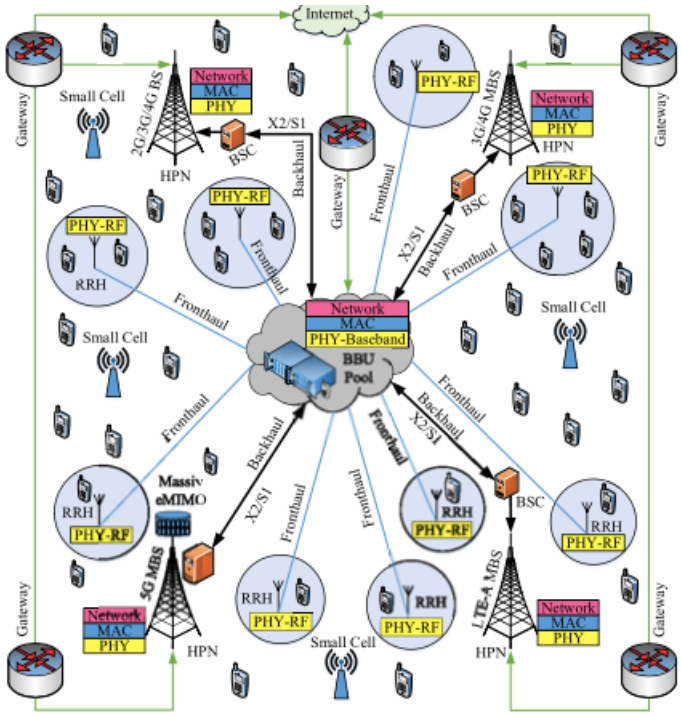


Figure 1.2: H-CRAN from [4]

1. **Cloud and Real-Time virtualized BBU pool**: The scattered BBUs are integrated into one single BBU pool which is connected also with (HPN)s to coordinate RRH and (HPN)s communications.

2. **Improved transport network** where all RRHs are connected to its corresponding BBUs located in the BBU pool. Both RRHs and BBUs are interconnected via low latency and high bandwidth fronthaul links such as optical fibre. BBUs are connected with the macro base station and high power nodes via S1 and X2 which are respectively the data and control interfaces.

3. **Heterogeneous composition**: In this kind of RAN macro base stations, small base stations, and RRHs can coexist. The macro base stations control the network, manage mobility, and improve performance; whereas small cells and RRHs increase system capacity and decrease transmission power. The radio functionalities are implemented in the RRHs whereas the upper layer ones are implemented in the BBU pool. The HPNs instead are equipped with all the network functionalities.

H-CRAN has been extensively studied by the researcher. As shown in [21] [23] [24] [25], H-CRAN has been evaluated from the energy and performance point of view. In

particular, in these works, the authors focused on both Dl and Ul data rates, and explored mechanisms to enhance the throughput and fulfil the QoS requirements. Also, they focused on the power allocation strategies to minimize the energy consumption of the H-CRAN. As in [26] , resource allocation is also a key aspect to consider. They analysed the state of the art and derived challenges to allow the evolution of such radio access networks. The authors, in [27] focused on the resource allocation strategies that have been designed to assign resource blocks to users to fulfil specific service-related requirements. In particular, they focused on three levels namely the spectrum, the infrastructure and the network. They provided resource allocation strategies introducing improved management of the radio resources. Additional example of improved management of radio resources introducing benefits in terms of system efficiency can be find in [28] and [29]. About CAPEX and OPEX expenditures, it is worth mentioning the work [11], where the authors proposed a theoretical framework to evaluate the CAPEX and OPEX of H-CRAN deployment. They have shown that such cloud ran based-deployment allows the telco operator to save up to 15% of the deployment cost of legacy 4Gn.

1.1.3 V-CRAN

Along with the evolution of radio access networks, another evolution process was occurring in the core network segment. The key enabling factors for the process in the core network were mainly Network Function Virtualization (NFV) and the Software Defined Network (SDN) approach. After extensive studies focusing on the application of these two factors on the core network, SDN and NFV have been extended even on the RAN. Numerous challenges came along with this extension because of wireless network features. However, the application of SDN and NFV on the RAN allowed the design of a new type of radio access network that follows the trend of the already presented RAN and it is known as **V-CRAN**. An example of V-CRAN architecture is presented in the fig.1.3. As described in [30] V-CRAN is composed by

- **Digital Unit Cloud (DU Cloud)** hosting all the commercial servers providing baseband processing and also the *layer 2* and *layer 3* related functions. All DUs are interconnected via *layer 2* switch also used for signalling among them.
- **Time-Wavelength-Division Multiplexed Passive Optical Network (TWDM-PON)** is composed of a single Optical Line Terminal (OLT) and multiple Optical Network Unit (ONU)s. The OLT is connected with the DU cloud providing each DU with an optical transceiver and a Line Card (LC). The LC conducts

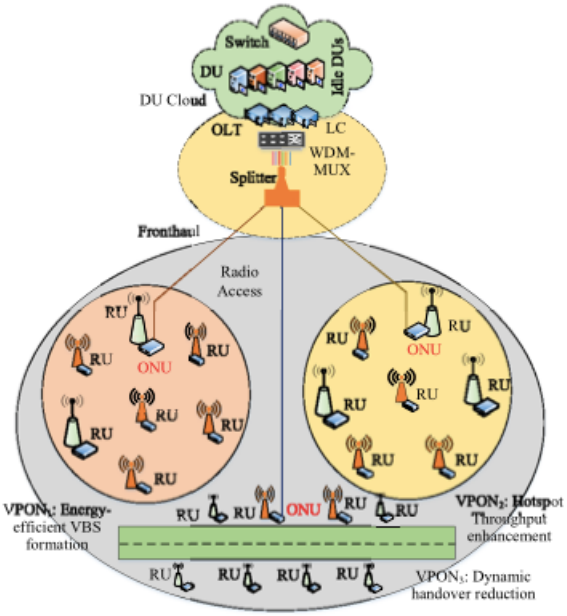


Figure 1.3: V-CRAN from [4]

optical-electrical conversion and delivers the traffic. All LCs of every single OLT are connected with a Wavelength-Division Multiplexed Multiplexer/De-Multiplexer (WDM-MUX) to separate traffic based on wavelength. An ONU is located remotely from DU Cloud and at the end of every optical channel aiming to enlarge the coverage of a TWDM-PON. It is co-located with an RU and is equipped with a re-configurable transceiver.

- **Fronthaul**: that is higher bandwidth and lower latency network used to exchange data and signalling between Radio Units (RUs) and DUs. To fulfil all these requirements, the utilization of optical channels over a single fibre is considered one of the viable candidates for V-CRAN's fronthaul. The fronthaul shares high-capacity optical channels among V-BSs using TWDM-PON technology.

- **Virtualized Base Station (V-BS)** it is a virtualized base station. The virtualization can be done at the hardware level, reducing the OPEX and increasing energy efficiency or flow level increasing the efficiency of resource multiplexing as described in [31]. The V-BS shares radio equipment at the hardware level and runs multiple protocol stacks of a BS in the form of software

- **Virtualized Passive Optical Network (VPON)** is defined as a set of ONUs sharing one or more wavelengths. The VPON can be also seen as a virtualized

channel that is used to connect users with the DU. Every DU in the DU Cloud is dedicated to a single VPON using LC. It specifically means that multiple RUs of a VPON are dynamically grouped and controlled by a single DU located in the DU Cloud. According to the operational condition of the network, the VPON can be dedicated or not.

Additional details related to this radio access configuration can be found in the document [32]. Even the V-CRAN has been extensively studied by researchers from a different perspective. An example of V-CRAN can be found in [33] where the authors present a virtualization framework of Base Station in the LTE network called ''OpeNB''. The proposed framework utilizes SDN, OpenFlow, and virtualization technology in the RAN and more specifically in the BS. Regarding performance, it is worth mentioning the following works. In [34], the authors propose a framework assigning limited rule space to maximize the number of social IoT groups contemporary satisfying latency in V-CRAN context. In [35], the authors analyse the performance of V-CRAN in terms of throughput, system stability and fairness considering two resource allocation strategies: *Dominant Resource Fairness* and *Proportional Fairness*. In [36], the authors design the macro-level and the micro-level metrics of virtualization of Base station (BS) to evaluate the performance of the virtualized base station. Even in V-CRAN, resource allocation plays a key role and it has been the subject of different studies. In [37], the authors propose an algorithm of resource negotiation for network virtualization in the context of a heterogeneous LTE-A network to achieve on-demand delivery of radio resources. In [38] instead, the authors investigate the resource virtualization, coordination, and its dynamical allocation by a hypervisor among different virtual operators. Also in [39], a framework for resource allocation is provided by the authors. Regarding energy consumption, in [39]the authors shows also that the KORA framework guarantees optimal performances to the negligible energy consumption cost of almost 7%. In [40], The authors claim that the proposed scheme reduces average power consumption by 65%, 6% and 3% less than for the Distributed Baseline, the First Fit Decreasing, the Heuristic Simulating Annealing algorithms respectively. In [41] and [40], CAPEX and OPEX expenditures has been analyzed. In particular, the cost analysis has been integrated into their used models optimizing either the energy efficiency or the system performance.

1.1.4 F-CRAN

It is expected that in the next years the volume of traffic and the device connectivity density in the 5G network will reach a magnitude that never has been seen. It is expected to have scenarios with billions of connected devices such as cars, sensors and so forth. This will have a huge impact on the performance of the radio access segments. To sustain such a huge increment, the cloud-based ran technology solution appears to be the most attractive. However, the previous solutions have limitations such as large end-to-end delays, traffic congestion, processing of huge amounts of data and expensive cost of deployment. Therefore, a new cloud-based architecture overcoming these limitations has been proposed by CISCO in [42] with the name of "FOG computing". The main idea is the extension of cloud computing near the user premises when you need it. As described also in [43], FOG computing allocates a large amount of processing, storage, communication, control, configuration, measurement, and management functions at the edge of the mobile network. The extension of this idea in the radio access segment led to the definition of a new cloud-based ran architecture also known as F-CRAN. As you can see in the fig.1.4, this new type of architecture is composed by *terminal layer*,*network access layer*,*cloud computing layer*.

- As shown in fig.1.4, the Fog Access Points (F-APs) in the network access layer and Fog User Equipments (F-UEs) in the terminal layer are considered as the mobile fog computing layer. In the terminal layer, F-UEs do also access the HPN to receive information related to system signalling. Moreover, the neighbouring F-UEs can communicate with each other in the terminal layer using D2D communication mode where an additional F-UE acts as a relay (F-UE5 and F-UE6 which are using F-UE7 as the relay).

- The network access layer is composed of HPNs and F-APs. The HPNs are deployed to provide system information related to signalling to all F-UEs in the area. However, the F-APs process and forward the data received from the F-UEs. F-APs and HPNs are interfaced with the BBU pool in the cloud computing layer through fronthaul and backhaul links, respectively.

- In the cloud computing layer, the BBU pool is compatible with that of H-CRAN.

F-CRAN has been studied from various dimensions such as performance, energy consumption, and resource allocation. In particular, the authors in [44] considers the

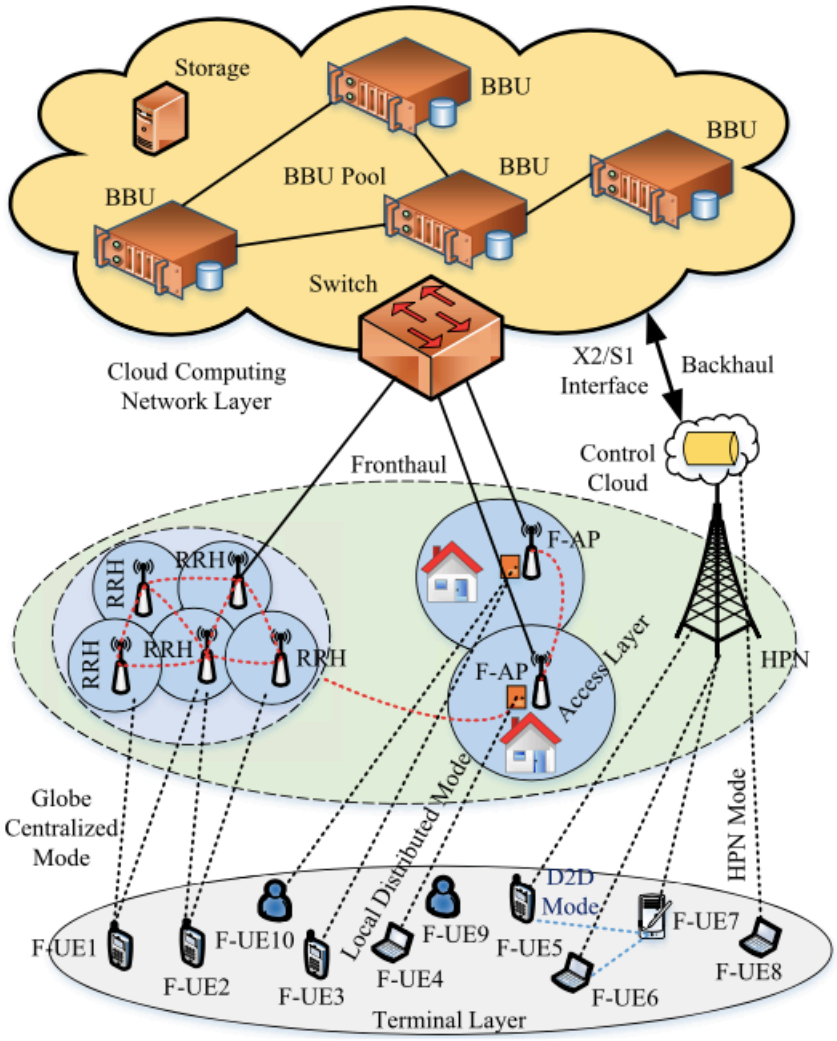


Figure 1.4: F-CRAN from [4]

F-CRAN with hierarchical content caching. Each F-AP is equipped with an individual cache and a part of requests can be responded to locally. The stochastic geometry is also used for deriving the average ergodic rate for the content transmission and queuing theory has been utilized to retrieve the waiting delay and latency ratio. In [45], the authors focused on the maximization of energy efficiency taking into account delay constraints and resource reuse constraints. In [46] instead, the authors focused on the minimization of the Dl latency in F-CRAN. Moreover, the authors in [47] study a latency-centric understanding of the degrees of freedom in high signal-to-noise ratio regime in the F-CRAN with limited available resources such as fronthaul capacity, cache storage sizes, power and bandwidth of the wireless channel. Also, the previous works, [48] [45] [47] and [49] can be considered as an example of

energy efficiency studies. About resource allocation, it is worth mentioning the following studies, [50] [51] [52] , where the authors investigate different resource allocation strategies. In particular, they focused on how to provide a good trade-off between communicating and computing resources in a time domain in F-CRAN scenario.

1.2 Network Function Virtualization (NFV) and Software Defined Network (SDN) paradigms

In this section, we describe the key enabling software technologies that are making possible the evolution process towards the 5G RAN networks. These software technologies are **SDN**, **NFV**, and **Network Slicing (NS)**. These technologies have been widely used to design the RAN that we presented in section 1.1. The **SDN** approach represents a new way of thinking about the networks. As depicted by the fig 1.5 the legacy systems, such as 4Gn, are vertically integrated. This means that

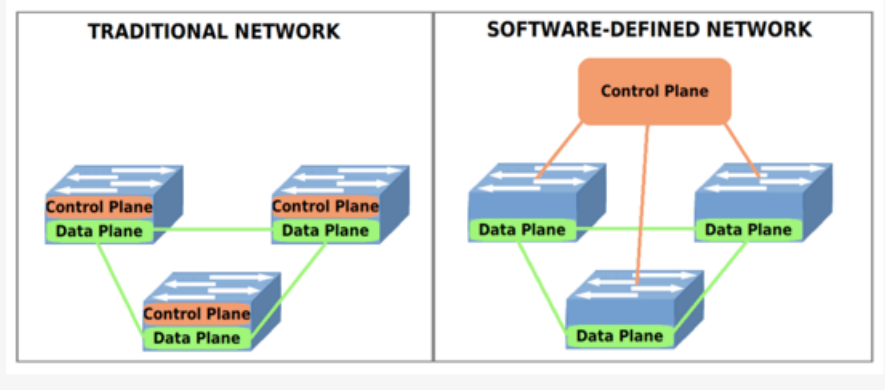


Figure 1.5: SDN based networks vs Traditional networks from [53]

the control plane and data plane are paired together. This integrating approach to traditional networks makes it difficult to manage and hard to configure them when the need for an upgrade comes up. Given this main limitation, an attractive solution appeared to be a separation of the control plane from the user plane. The SDN architecture is depicted in fig. 1.6. It consists of three layers: *infrastructure layer*, *control layer* and *application layer*. The *infrastructure layer* is the bottom one and it is composed of all network devices and hardware. As shown in the architecture, the *infrastructure layer* interacts with *control layer* via control plane interfaces, also known as southbound interfaces. The *control layer* can be composed of multiple SDN controllers. All the network logic is centralized in this layer and possibly distributed among these multiple controllers. Also, these can communicate by specific interfaces

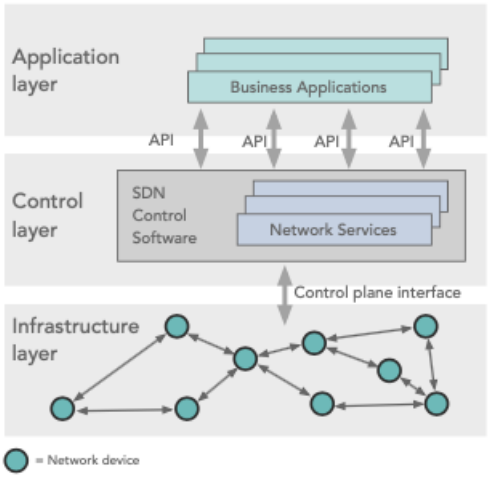


Figure 1.6: SDN architecture

that in literature are known as east and ovest bounds. The *application layer* then is the upper one, where service providers, network operators and application developers operate to fulfil their business demands such as bandwidth, traffic, access control, QoS, energy usage and so forth. The 5G mobile network infrastructure is based on SDN. This will help operators to provide efficient communication between applications and services in the cloud and end-users. The deployment of SDN in 5G mobile network increases efficiency in the allocation of radio resources through centralization and seamless mobility over diverse radio access technologies. In literature, there are many studies covering the application of SDN onto the 5GRAN. For example, the [54] can be considered the early application of software-defined networking on the radio access segment. Here, the authors model an architecture where the SDN controller adjusts bandwidth for each radio Access Point Access Points (AP) and BBU dynamically. The deployed SDN controller has the responsibility of management and selection of routes for all RAN and core network connections. In [55], the author present a radio architecture leveraging the SDN approach called "Soft-Air". In [56], instead, the authors propose a multi-tiered cloud controller scheme and event processing mechanism for the Software-Defined Wireless Network (SDWN) architecture of 5G mobile communication. Also in the paper [57], the authors present the architecture of an SDN-supported system, OpenFlow standard, and alternatives for the deployment of SDN-based protocols and services. The **NFV** architectural framework is illustrated in fig. 1.7, which is proposed by the ETSI [58]. The ETSI NFV framework consists of three parts: the Network Functions Virtualization Infrastructure (NFVI), Virtualized Network Function (VNF), and NFV Management and

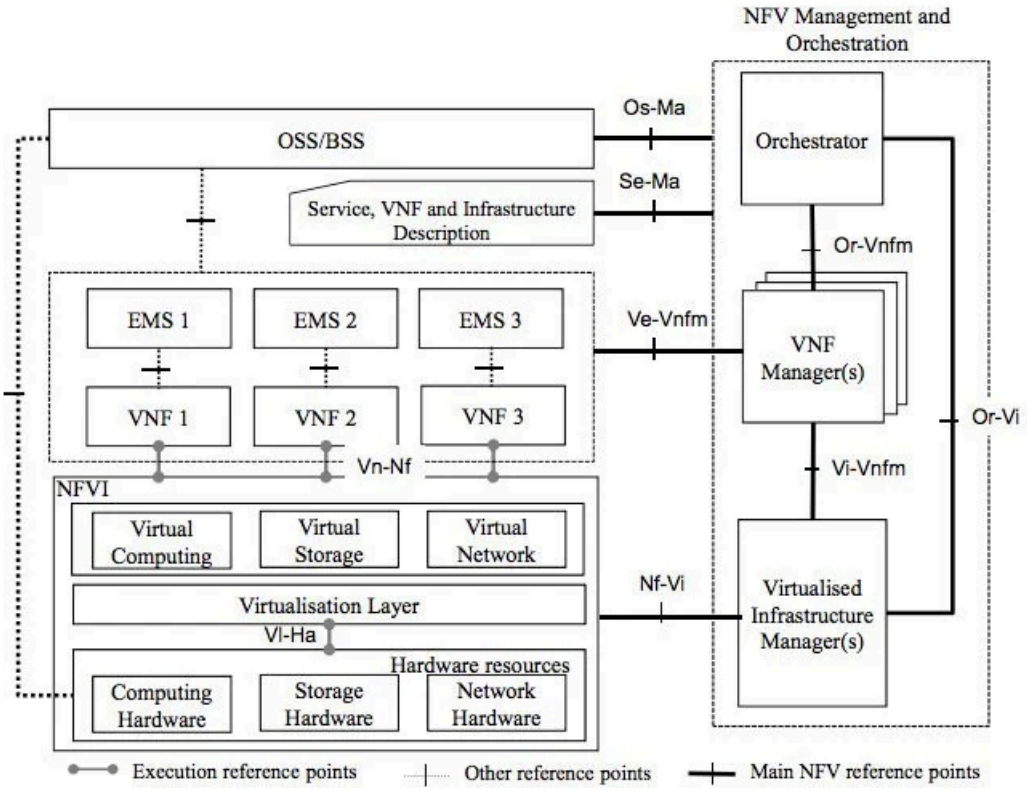


Figure 1.7: NFV architecture from [58]

Orchestration (NFV MO). The NFVI corresponds to the data plane providing virtual resources to the VNFs to be executed. The VNF is the software implementation of a network function and can run over the NFVI. The VNF corresponds to the application plane and consists of connected VNFs, that, in this context, can be considered applications as well. The NFV MO part of the architecture corresponds to the control plane. It is responsible for the orchestration and life-cycle management of hardware and software network resources, which are used to support the virtualization-based process. Moreover, the management and the orchestration entity have the responsibility to build connection among different VNFs and interacts with the OSS/BSS landscape. This component allows the ETSI framework to interact with existing network management. The NFV system is driven by a set of metadata consisting of the requirements for service, VNFs, and infrastructure allowing the NFV MO to properly deliver the required service. Considering the 5G RAN architecture, most of the user plane and control plane functions can be virtualized. The virtualization of functions in 5G RAN introduces benefits in several network aspects such as energy consumption and efficiency in resource allocation. According to their architectures, SDN and NFV can be integrated among themselves. This is because they have highly com-

plementary features, as shown in [59], [60], [61]. Also in [62] the author, investigate the deployment of NFV in software-defined NFV architecture. The authors present software-defined NFV architecture and describe also the relationship between SDN and NFV. In [63], the authors survey recent papers related to the deployment of NFV along with SDN to highlight the advantages and disadvantages of both approaches.

1.3 Network Slicing on Radio Access Network

The adoption of SDN and NFV created the landscape for a breakthrough technology known as Network Slicing NS. The main idea of NS is depicted in the fig.1.8. As you can see, the NS allows the telco-operator to partition the network in a structured, elastic, scalable, programmable and automated manner. Those network partitions have a different level of isolation. Also, they have their specific characteristics and will be considered as a single logical entity of the same physical shared network. Each use-case/business scenario can exploit this network entity according to the needs. Be-

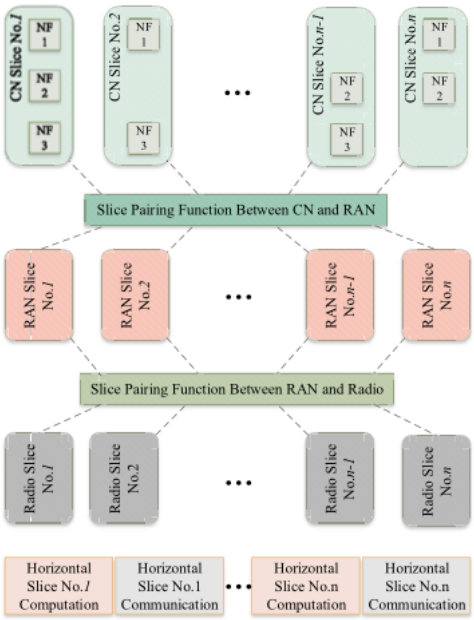


Figure 1.8: NSidea from [4]

fore going deep into the details, it is worth noting that NS has implied a re-engineering of both the core network and radio access network. The core network adopted the software architecture paradigm known as Service Oriented Architecture (SoA). SoA allows telco-operator to avoid disruptive changes in the legacy systems. This means

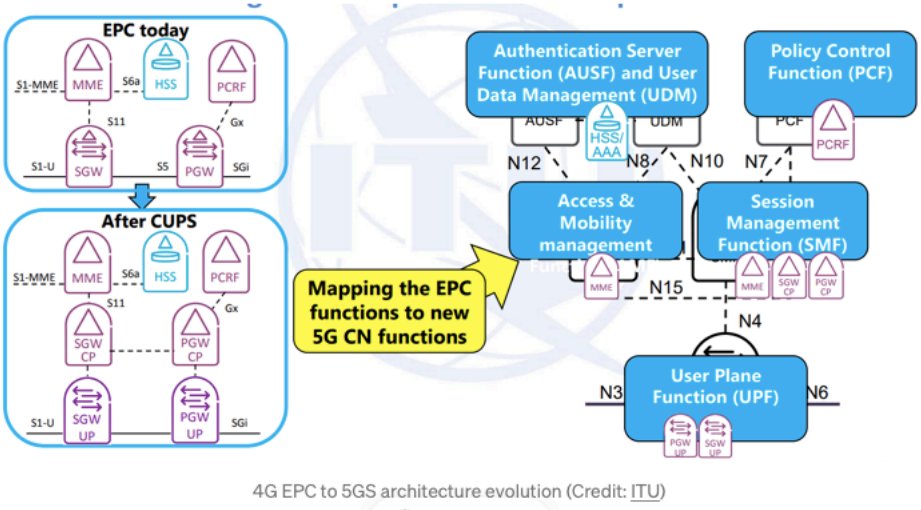


Figure 1.9: From 4G to 5G core network

a cheaper converting process from 4Gn towards 5Gn. In fig.1.9 you can see the no-disruptive converting process that from 4Gn ended up in 5Gn. As you can see the first step was the separation of the control plane and user plane functionalities according to the SDN approach. The second step was the re-arrangement of those obtained functionalities from the previous step in services according to the SoA paradigm. After that, following the NFV paradigm, those services can be considered a pool of network functions that can be instantiated and chained together to provide the needed service. About the RAN instead, the converting process produced a new architecture that is presented in 1.10. Such architecture corresponds to the straightforward application of the SDN approach on legacy radio components. Some of them have been removed and others have been specifically designed. Also, the migration from the 4G radio access network has been characterised by the following prototyping new generation of radio access network (NG-RAN) architectures cornerstones that are:

- non-standalone (NSA) NG-RAN
- stand alone (SA) NG-RAN

(NSA) NG-RAN can be considered the intermediate architecture integrating the New radio configuration and the legacy network system of LTE. Instead, the (SA) NG-RAN architecture prototype consists in a set of connected 5G base station (gNB)s, with a 5G core network through a set of logical interfaces. Also, these gNBs can be interconnected through the Xn interface to improve mobility and management functions.

The gNB's functionalities are sometimes distributed. In that case, the resulting architecture is formed by a central unit 5G base station controlling unit (gNB-CU) controlling one or more distributed units 5G base station distributed unit (gNB-DU) through the F1 interface. A distributed unit is connected to a remote radio head RRH, i.e., the actual radio transceiver. Moreover, according to the SDN approach that introduced the separation of functionalities, the central unit can be split into two parts, one for control plane functions gNB-CU-CP and one for user plane functions gNB-CU-UP. Interested readers can find further details about this architecture in [64]. However, going back to the fig.1.8 showing the system architecture of network slicing, also presented in [65], is possible to observe that technically it is composed by Core Network (CN) slices, RAN slices and radio slices. Each slice in CN is built with a composition of chained Network Functions (NF) together. The network functions can be shared among multiple slices or they are tailored to a specific slice. Generally, there are two slices at least plus pairing functions connecting all of them. The first pairing function is between CN slices and RAN slices, and the second pairing function is between RAN slices and radio slices. The paring function routes communication between the radio slice and its appropriate CN slice. The pairing function between RAN and CN slices can be static or semi-dynamic configurations to achieve the required network function and communication. Network slicing has been widely studied. For example, in [66] the authors studied the deployment of the network slice from the profit point of view. In this work, the main purpose is the optimization of profit modelling of traditional telecommunication networks by exploiting the knowledge of the network slice. Moreover in [67] the authors designed an algorithm that allocates requests of network slices, which maximizes the total revenue of the network. A comprehensive work accounting for different aspects related to the network slice is [68]. In particular, they review the state of the art restricted to the network slice with a particular interest in resource allocation, virtualization technologies, orchestration process and isolation function. Also, they provide use cases that can be handled with NS and then they list the existing challenges related to NS deployment.

1.4 New Promising Massive MiMo System: The Cell-Free Massive Mimo

In this section, we introduce the recent and appealing massive MiMo system also known as Cell-free massive mimo (Cf-maMiMo). It is expected that this technological breakthrough copes with these new demands of QoS related to the deployed services

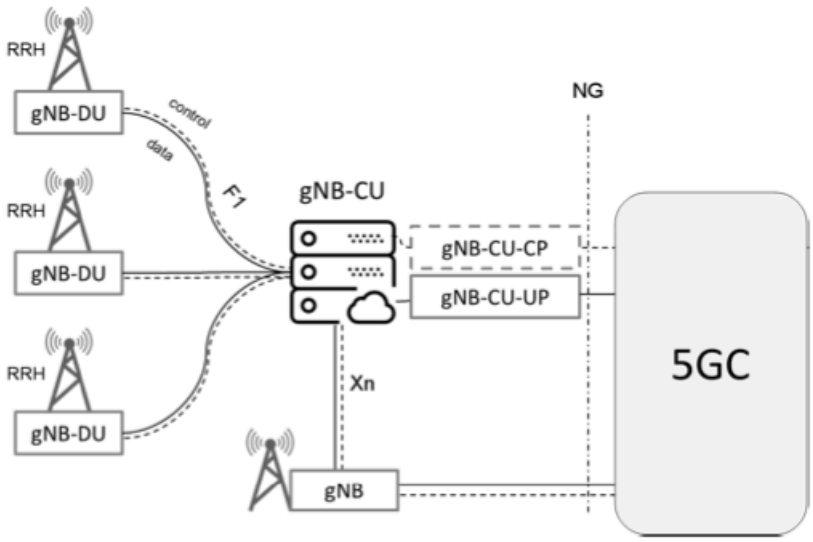


Figure 1.10: Overall architecture NG-RAN

in 5Gns. In Cf-maMiMo, a large number of distributed access points are connected to a central processing unit and serve a smaller number of users over the same time-frequency resources. This system can be flexibly integrated with various emerging techniques and technologies for 5G to benefit from improved performance from different perspectives. Despite the substantial reported theoretical gains of Cf-maMiMo systems, the full picture of practical scalable deployment of the system is not clear yet. As described in [69], the conventional embodiment of massive MiMo systems is the co-located ones in which macro BSs are equipped with a massive number of antennas communicating simultaneously with a smaller number of users at the same time-frequency resource through beamforming. According to the works [70] [71] [72], the large system dimensions lead to favourable propagation where the channels of different users are asymptotically orthogonal. Hence, simple linear pre-coding and detection techniques can be employed at (BS)s for downlink (DL) data transmission and Ul data detection, respectively. Also, It has been shown that the deployment of co-located massive MiMo systems can significantly deliver high data rates as well as enhanced link reliability, coverage, and/or energy efficiency. Nevertheless, the system suffers from performance degradation for cell-edge users due to the lower channel gain from the serving cell as well as the initiated interference from neighbouring cells. The side effect of co-located massive MiMo deployment systems deployment is the distributed massive MiMo. As presented also in [18], this MiMo system has a large number of distributed single or multiple-antennas-equipped APs that are connected

with a central process unit by high-speed fibre or wireless backhaul/fronthaul links. This kind of MiMo system has been widely studied, as in [73] [74], and results show that it outperforms co-located system in terms of better data rate values, also it provides a good level of QoS even for those users that are located at the cell-edge thanks to the high provided diversity gain. This is because each user at the edge receives large-scale components from different APs. Recently, an alternative to the distributed massive MiMo system has been proposed. It is known in the research and the industrial environment as Cf-maMiMo. As you can see from the fig.1.11 and also

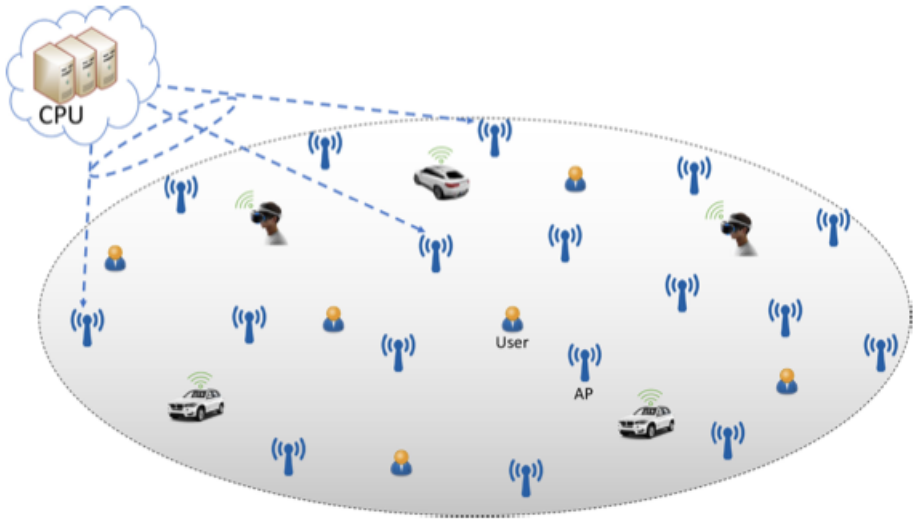


Figure 1.11: Overall architecture of Cf-maMiMo from [75]

presented in the seminal work [76] of Marzetta et. all, in the Cf-maMiMo system all APs are connected to a CPU. It allows them to serve all users over the same time-frequency resources through applying spatial multiplexing techniques e.g [77]. The performance of Cf-maMiMo system is compared with that of small-cell systems in terms of the achievable Dl and Ul data rates [78]. As in [79] [80] the Cf-maMiMo outperforms its competitor and can provide a uniformly good service for all users within the coverage area with a tailored power allocation mechanism. Due to its performance, the Cf-maMiMo operation is strongly recommended for beyond-5G networks due to its ability to assure seamless mobility support without overhead due to handovers. This in turn will provide improved QoS fulfilling the 5G expected requirements. As presented in [81], the Cf-maMiMo combines the best aspects of ultra-dense cellular networks with the cellular massive MiMo technology to overcome their respective weaknesses. In particular, due to the main properties which are many geographically distributed APs whose coverage area is not divided into disjoint cells

and each UE is served by all the surrounding APs, the massive MiMo processing resolves the interference situation that limits conventional ultra-dense networks and leads to a network free from cells. Moreover, the large number of distributed AP antennas allows for mitigating the large SNR variations that usually limit the efficiency of conventional cellular massive MiMo. As described in the [76] the original motivation behind Cf-maMiMo was to design a new network infrastructure capable of providing uniform area throughput. A Cf-maMiMo system can also be viewed as a user-centric network because each UE will only be influenced by the signals from the closest surrounding APs. Such organization can exploit the flexibility of the presented RAN architecture in the section 1.1, such CRAN, H-CRAN and so forth.

1.5 Summary

In this chapter, we introduced the state of the art of the 5G-based radio access networks. In particular, we emphasised their architectures and how they evolved from the early version, represented by the Cloudyfied radio access network, also known as C-RAN, towards the more complex, represented by the FOG-CRAN. A common trend in this architecture is represented by the need for flexibility. Then, we shifted our focus to the new networking paradigms, SDN and NFV, which are considered the main enabling software mechanisms supporting the deployment of those architectures. The extreme integration of both provides what today is known as Network Slicing. This represents the main technological breakthrough for the introduction of radio access network flexibility from the software point of view. Last, but not least, we introduced the concept of the Cell-free massive mimo that recently has been recognised as one of the radio access network configurations providing topology flexibility.

Chapter 2

Bandwidth Part Management

During the evolution process toward a 5G radio access network, the main issues were the lack of flexibility and programmability. A most promising solution to overcome the problem of the lack of programmability is the adoption of software-defined networking approaches in radio access networks. This means that, in this network segment, it is expected the deployment of an SDN-controller leveraging the new radio technologies eg. Flexible numerology, and Bandwidth part. In other terms, it is expected that the new SDN controllers are equipped with core functionalities able to sustain the Bandwidth Part Management activity including the underlying concept of flexible numerology and bandwidth part. Due to this consideration, this chapter is arranged as follows: In sec: **2.1 Flexible numerology**, we introduce the concept of **Flexible numerolgy (Fn)** which represents the essential fundamental for the definition of bandwidth part. This last one is explained, instead, in the sec: **2.2 Definition of the Bandwidth Part**. The sec: **2.3 The Bandwidth part management**, we explain why dynamic bandwidth part management plays a vital role in the evolution process toward 5G network. In the sec: **2.4Dynamic selection of numerology and bandwidth part for effective ORAN slicing**, we introduce the realised work on bandwidth part management showing that a dynamic and SDN-assisted version of it performs better than the static one.

2.1 Flexible Numerology

The Long Term Evolution (LTE) waveform has a fixed structure that is optimized for applications requiring a high data rate. However, there exist others applications that cannot be supported due to the inflexibility of the waveform. The technical specification [82] allows us to understand the limited flexibility of the LTE. In par-

ticular, the extended cyclic prefix configuration which is utilized by macro-cell BSs at all times to keep the system operating at larger delay spreads, introduces reduced spectral efficiency. The The 5G eco-system has been designed for providing a wide variety of services that have to be contemporarily deployed. As described in [83], this can be done by introducing flexibility in waveform parameters. This flexibility gives us, in turn, the chance to make Enhanced Mobile Broadband (Embb) experience possible everywhere, including highly mobile User Equipment (UE), connected to macro-cells. Also, it enables reduced latency values and improved reliability transmission, also known as ultra-reliable and low latency communications Ultra reliable low latency communication (Urllc) and massive machine type communicationsmachine to Machine Type Communication (mMTC) for suitable scenarios e.g smart cities or high-density connection-based scenarios. To achieve this flexibility feature, a new degree of freedom in waveform designing progress has been introduced and it is known as flexible numerology. This one takes over the fixed numerology characterizing the legacy LTE system. As described in [82], NR allows simultaneous multi-numerology utilization. Flexible numerology has been widely studied. For example, the authors in [84] focus on the design of multi-numerology systems. Specifically, the authors devise a framework that provides simultaneously numerous services in a unified frame radio and their results show the achievements of better frequency spread immunity and spectral efficiency by exploiting the doubly dispersive channel characteristics of the users. Multi-numerology systems with similar performance and results have been also studied in [85], [86], [87]. Due to the specification [82], it holds that higher numerology indexes correspond to larger Sub Carrier Spacing, ranging from 15 and 480 kHz by following the equation:

$$\Delta f = 15 \times 2^{\mu} \quad KHz \tag{2.1}$$

The numerology μ depends on different factors (i.e., service requirements, deployment type, carrier frequency, etc.). It allows the definition of different sub-carrier spaces mitigating inter-carrier interference and phase noise at mm-Wave frequencies. As sub-carrier space widens, the TTI assumes smaller values ranging from 1 ms to 31.25 μs. Other parameters also change from numerology to numerology. In fig. 2.1 are summarised all the numerologies which have been standardised by 3GPP and it shows the different sub-carrier spacing that we can use to design the proper waveform according to the needs. Moreover, based on different numerology we have different time slot lengths, fig. 2.2, and per slot number of symbols. Also, according to numerology and type of cycle prefix (normal or extended), we can have a different

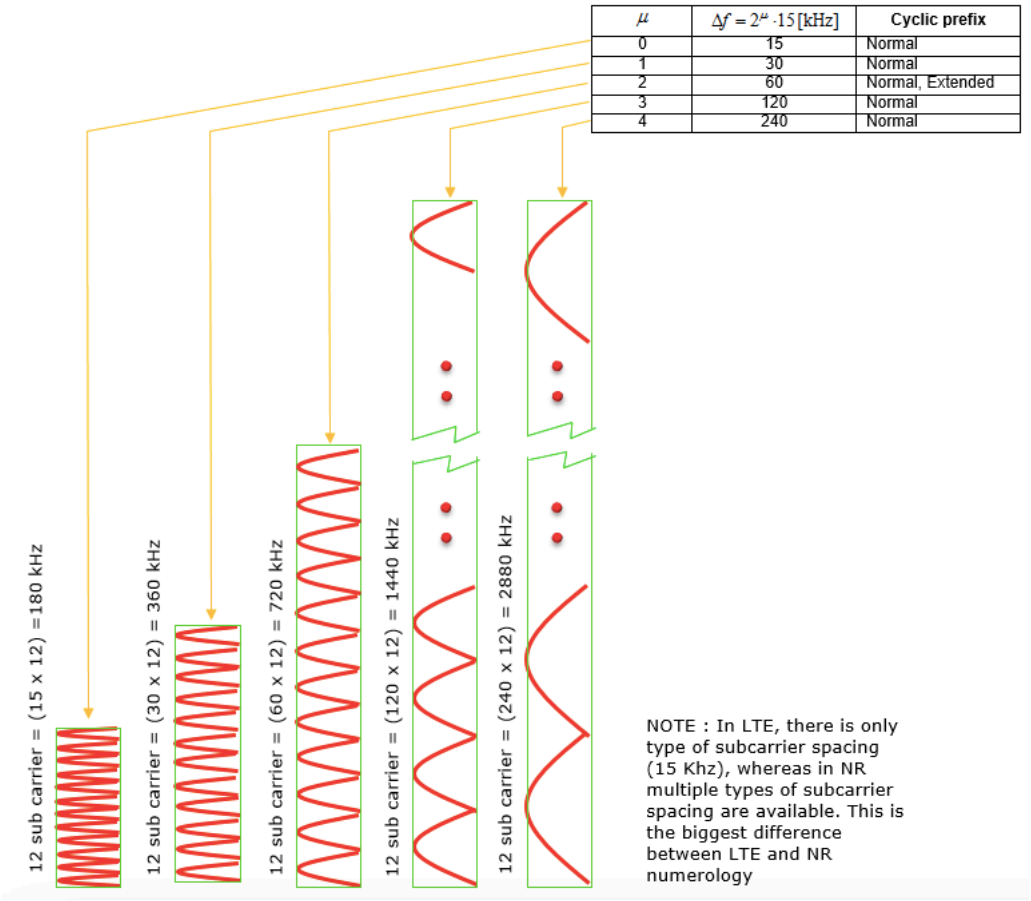
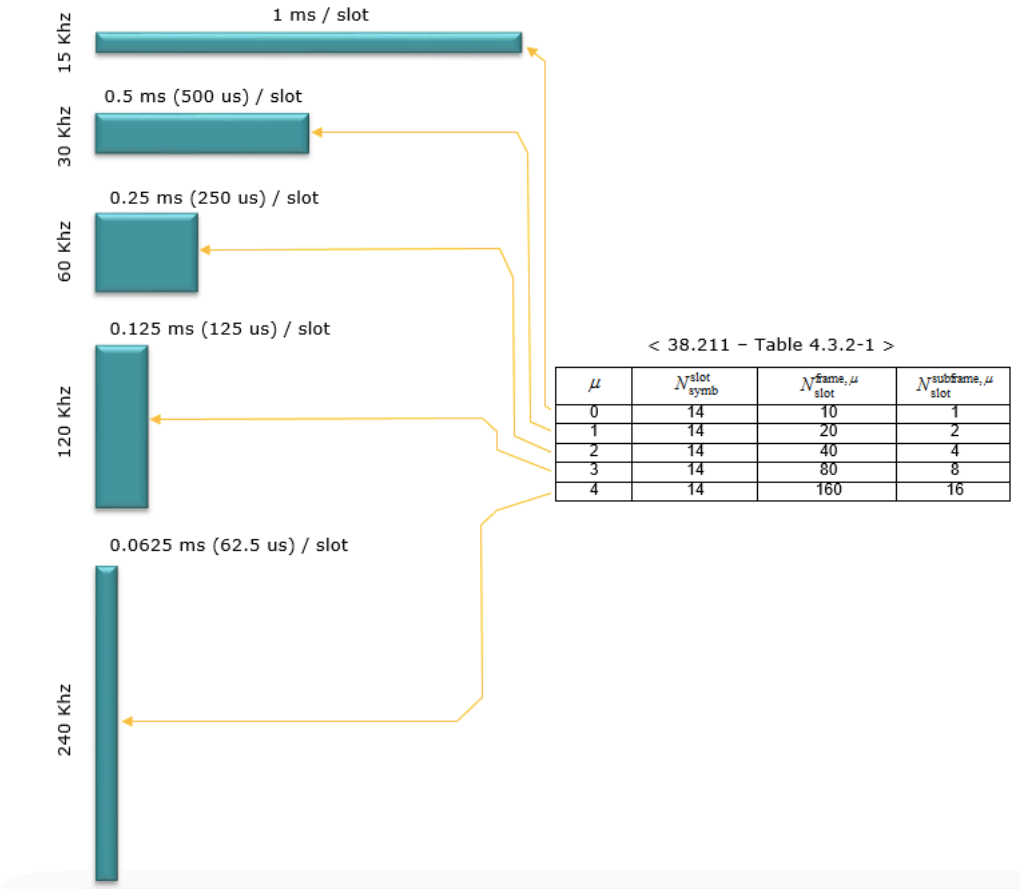


Figure 2.1: Flexible numerology from [88]

arrangement of the radio frame. Despite the multiple chosen numerology, the length of the radio frame and the length of the sub-frame remain the same, that is 10 and 1 ms respectively. Fig. 2.3 summarises all the possible radio frame configurations. Here, we can see how the numerology affects the duration of time slots in the time sub-frame and also how it affects the dimensions of the resource grid. Moreover is possible to observe how the physical resource block is arranged.

2.2 Definition of the Bandwidth Part

The flexible numerology enabled a breakthrough radio technology that is called Bandwidth Part (BWP). As described in [90] one of the reasons for introducing BWP is to sustain the UE bandwidth adaptation to reduce device power consumption. The idea behind this is that UE may use a wide bandwidth when a large amount of data is scheduled while being active on a narrow bandwidth for the remaining time. Moreover, another reason behind their introduction is to support the heterogeneous UE devices' capa-



< 38.211 – Table 4.3.2-1 >

μ	N^{slot}_{symb}	$N^{frame,\mu}_{slot}$	$N^{subframe,\mu}_{slot}$
0	14	10	1
1	14	20	2
2	14	40	4
3	14	80	8
4	14	160	16

Figure 2.2: time slot duration following the chosen numerology from [88]

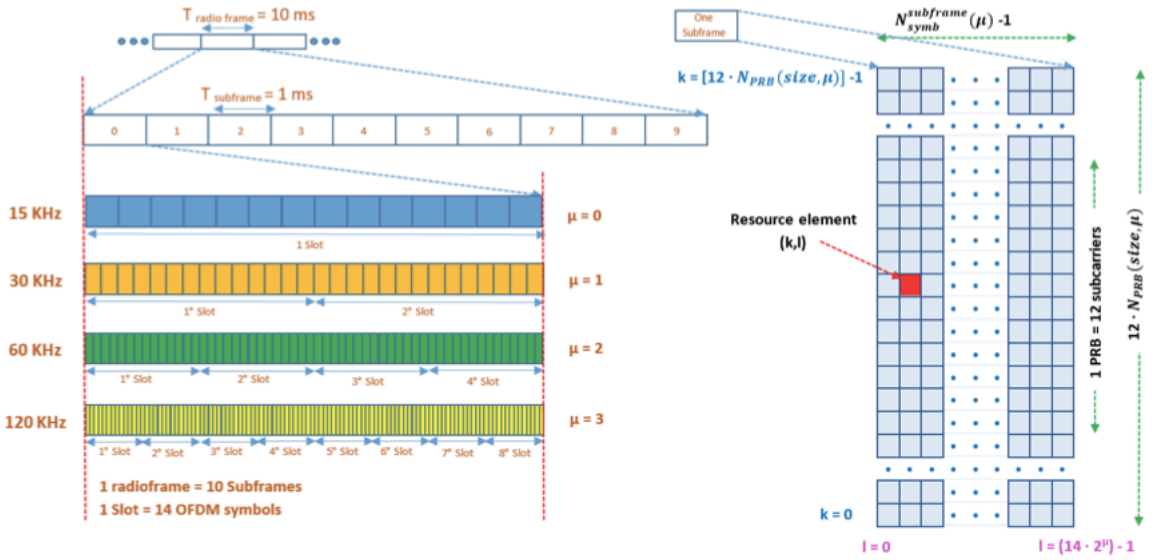


Figure 2.3: Frame configurations and resource grid from [89]

bilities. As described in [91], usually a base station may support a very wide channel bandwidth which may not be supported by some UEs. In this context, the BWP provides a mechanism to flexibly assign radio resources such that the signals for a UE are confined in a portion of BS channel bandwidth that the UE can support. At a high level, the BWP is a set of contiguous resource blocks which are configured within a channel bandwidth. According to the fig.2.4 New radio defines frequency ranges that

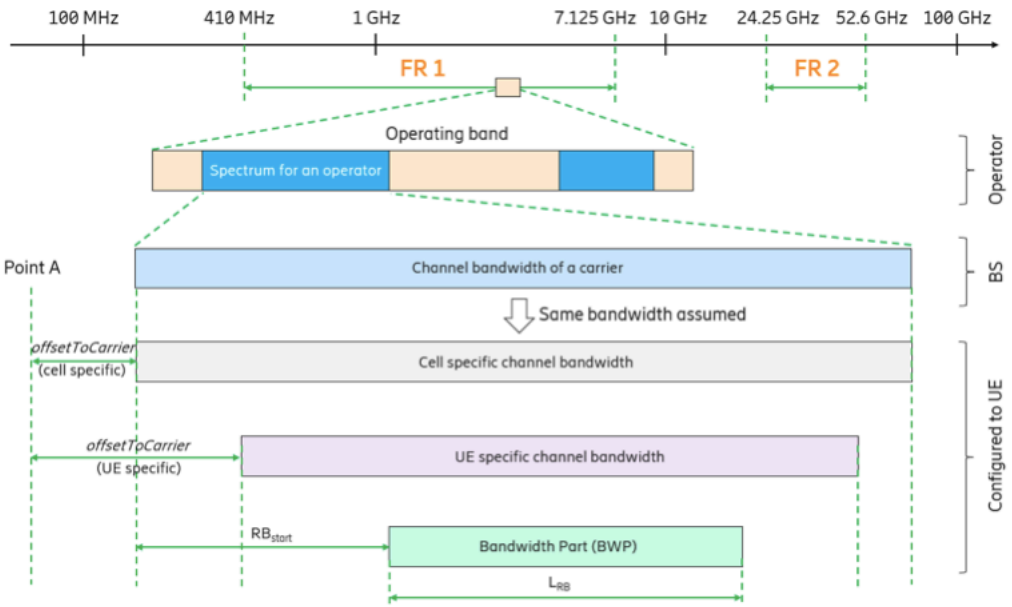


Figure 2.4: Introduction to BWP from [92]

are FR1 and FR2 ranging from 410 – 7125 MHz and 24.25 – 52.6 GHz respectively. In each FR, 3GPP further defines operating bands which are frequency bands associated with a certain set of radio frequency requirements. Within operating bands, operators may have different amounts of spectrum allowing them to define several frequency bands according to the required services. Their range goes from several MHz to a few GHz. Moreover, NR supports a range of channel bandwidths from 5 – 400 MHz which are also referred to as the bandwidth of an NR carrier. These NR carriers are used to accommodate diverse spectrum scenarios while limiting implementation complexity. The number of resource blocks Resource Block (RB)s that may be configured in channel bandwidth, shall meet the specified minimum guard-band requirements. An RB consists of 12 consecutive subcarriers in the frequency domain. Starting from a common entry point known as "Point A", specifically a BWP **consists of a set of contiguous RBs with given numerology on a given carrier that belongs to a common set of RBs**. For each serving cell of a UE, the network configures at

least one downlink (DL) BWP. It can be configured four BWP at most. The only constraint is that at most one Dl BWP can be active at a given time. Similarly, It can be done for the Ul configuration. Specific BWPs can be activated in a given time:

- **Initial up/dl link BWP**: these BWPs are used for initial access before a radio resource control connection is established. They are indexed with zero and are also referred to as BWPs 0. During the initial access, the UE performs a cell search based on a synchronisation signal block that is composed of a primary synchronisation signal, secondary synchronisation signal, and a physical broadcast channel. Also, the UE needs to read the System Information Block 1 carrying important information related to the initial ul/dl BWP configuration. Block 1 is transmitted on the PDSCH, which is scheduled by Dling control information on the PDCCH using the control resource set with index zero. Before the UE reads block 1, the UE's initial dL-BWP has the same frequency range and numerology as those of the control resource set. After its reading, the UE follows the initial ul/dl configuration contained in it and uses them to carry out the random-access procedure to request the setup of a radio resource control connection. the procedure describing the initial access of the UE in the system is provided in [93] and [94].

- **default BWP** For a serving cell, the network may configure the UE with a BWP inactivity timer indicating, for example, that the UE has no scheduled transmission and reception for a while on the currently active BWP. This offers the chance for the UE to switch its active BWP to the default one for power-saving reasons. This BWP can be configured, otherwise uses the same that it has already configured as initial dl-BWP.

- **first active BWP** The first active DL and UL BWPs are the active DL and UL BWPs upon RRC (re-)configuration for a Special Cell or activation of a Secondary Cell. In particular, according to the memberships of the cell to the master or secondary group, the special cell can be the one where the UE performs the connection or re-establishment procedure(master) or performs random access for radio resource control re-configuration(secondary).

The BWPs have configuration options that are also known as *common* and *dedicated* parameters. The first one is cell-specific, while the seconds are UE-specific. Given a specific dl-BWP with a non-zero index, the common parameters are *frequency domain location*, *bandwidth*, *Sub Carrier Spacing*, and *Cyclic Prefix*. The dl-BWP *dedicated*

parameters are *PDCCH*, *PDSCH*, *semi-persistent scheduling*, and *radio link monitoring configurations*. For the ul-BWP with non zero indexes instead, the set of *common parameters* includes those related to *random access, PUCCH, and PUSCH*. The *dedicated* ones include *PUCCH, PUSCH, SRS, configured grant, and beam failure recovery configurations*. For BWP with index 0 the possible configuration options are:

1. **Option 1** with cell-specific parameters limiting the functionality of the specific BWP. Usually, this kind of option is used to configure BWPs that have to be temporarily used, for example during the initial-access procedure.

2. **Option 2** with both cell-specific and UE-specific allow the full-featured BWP. This configuration is appealing in deployments where multiple ul/dl-BWPs are not needed. In this case, the network can set up a fully operational connection with a UE by only configuring this BWP with index zero using this option.

Moreover, New Radio supports configurations of up to four "RRC-configured" ul/dL-BWPs. The ul/dL-BWP with index zero following the Opt 1 only has cell-specific parameters and is not counted as an "RRC- configured"BWP. Therefore, additional **four** ul/dL-BWPs (with different indexes) may be consecutively configured. The ul/dL-BWP with index zero following the Opt 2 has both cell-specific and UE- specific parameters and thus is counted as an "RRC-configured" BWP. Therefore, additional **three** ul/dL-BWP may be consecutively configured. Another important aspect of

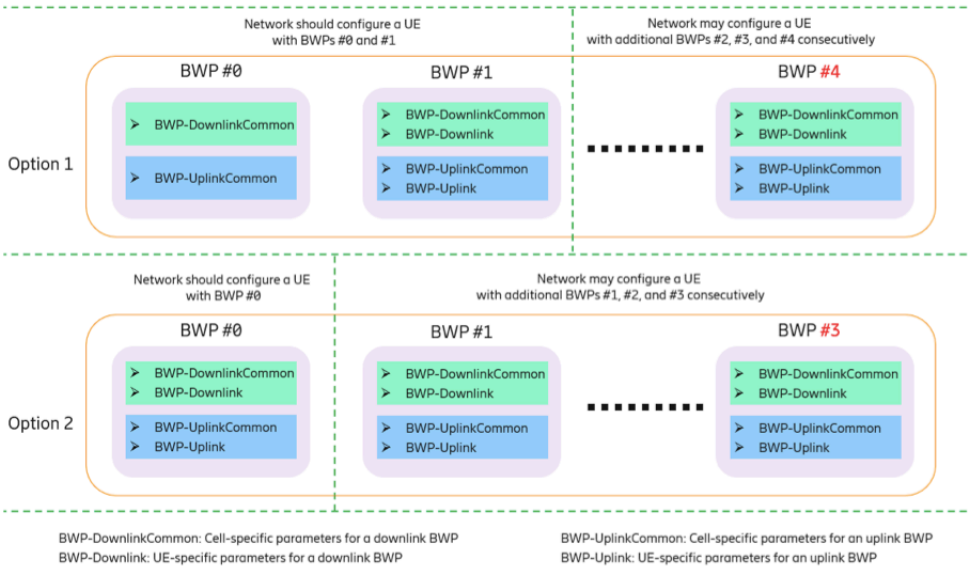


Figure 2.5: BWP configuration options from [92]

BWP is the switching operation. As we said, at a given time, only one BWP can

be active in both ul/dl links. In Frequency Division Domain (FDD), Dl and Ul can independently switch BWP but for time division domain (TDD), both downlink and uplink should simultaneously switch BWP. The BWP switching can be :

- **RRC signaling based**: The network can impose BWP switching using RRC re/configuration. This means that, upon receiving RRC re/configuration of active dl-BWP - and/or ul-BWP for serving primary cell, the UE activates that dl-BWP and/or ul-BWP properly indicated by respectively IDs. For a secondary cell, the UE doesn't activate the dl-BWP and/or ul-BWP immediately after receiving the RRC (re-)configuration, instead, the activation of the corresponding BWP is done at the time of secondary cell activation. If the network doesn't want to impose a BWP switch, the IDs of BWPs; are not declared and inserted in the RRC (re-) configuration message.

- **DCI based**: in 5G NR, DCI format 0_1 (uplink grant) and DCI format 1_1 (downlink assignment) are used to indicate BWP to be used. The field Bandwidth Part Indicator relates to the BWP in which the frequency resources provided via this DCI are located. This field is configured and can take 1 or 2 bits depending upon the number of UL/DL BWPs configured by RRC, excluding the initial ul/dL BWP. If the bandwidth part indicator field in DCI format 0_1 indicates a UL BWP different from the currently active UL BWP, the UE shall set the active UL BWP to the UL BWP indicated by this field in the DCI format 0_1. If the bandwidth part indicator field in DCI format 1_1 indicates a DL BWP different from the currently active DL BWP, the UE shall set the active DL BWP to the DL BWP indicated by this field in the DCI format 1_1. If a UE does not support active BWP change via DCI, the UE ignores the bit field Bandwidth Part Indicator;

- **the MAC entity itself upon initiation of Random Access procedure**: this kind of BWP switching operation is described by the fig.2.6; The MAC entity, upon initiation of an RA procedure on a serving cell, may perform BWP switching depending on whether or not PRACH occasions are configured for the active UL BWP. For the uplink, as PRACH occasions are **not configured** for active UL BWP, the MAC layer switches the active UL BWP to the initial uplink BWP. For the downlink, the switching will happen only if the serving cell is SpCell; in this case, the active DL BWP should also be switched to the initial downlink BWP. Instead of uplink, as PRACH occasions **are configured**

for active UL BWP, no need to switch active UL BWP. For the downlink, switching will happen only if the serving cell is SpCell and only if the BWP-id of active DL BWP is different from the BWP-id of active UL BWP; in this case, the active DL BWP should also be switched to the same BWP-id as the active UL BWP. The ultimate goal here is to bring both active UL BWP and active DL BWP onto the same BWP-id. After performing appropriate BWP switching (Case1 or Case2), the UE performs the RA procedure on the activated DL BWP of the SpCell and the activated UL BWP of the serving cell on which the RA procedure is triggered.

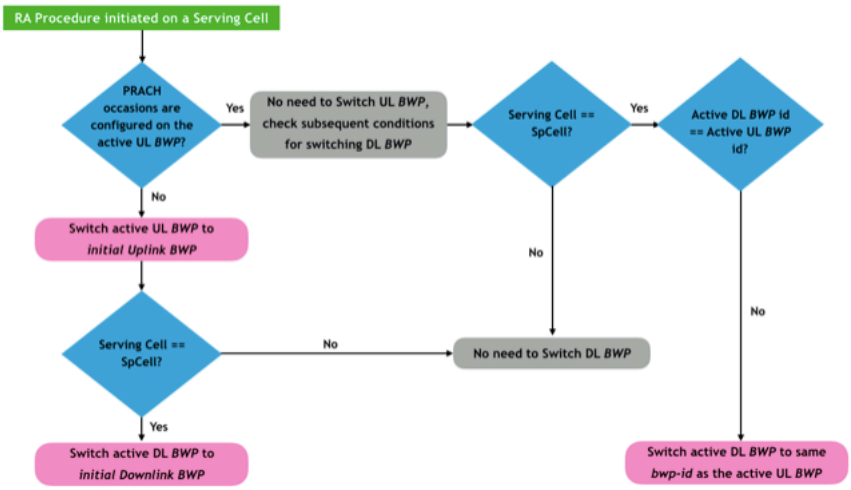


Figure 2.6: Mac entity based BWP switching operation from [95]

- **Timer based**: The network may configure an inactivity timer (BWP-InactivityTimer), which is used to switch the active downlink BWP after an amount of inactivity time specified by the timer field BWP-InactivityTimer. The expiry of the inactivity timer associated with a cell switches the active BWP to a default BWP configured by the network (if configured). If the default downlink BWP is not configured, the switching happens to the initial downlink BWP. In fig.2.7 is described the inactivity timer-based switching operation.

2.3 The Bandwidth part management

The SDN approach paved the way toward a more flexible and programmable radio access network. As we can see in [96] and [97], from the architectural perspective several early-stage approaches have been introduced with no real and practical implementation. In the next years, in the state of the arts, SDN-enabled RAN architectures

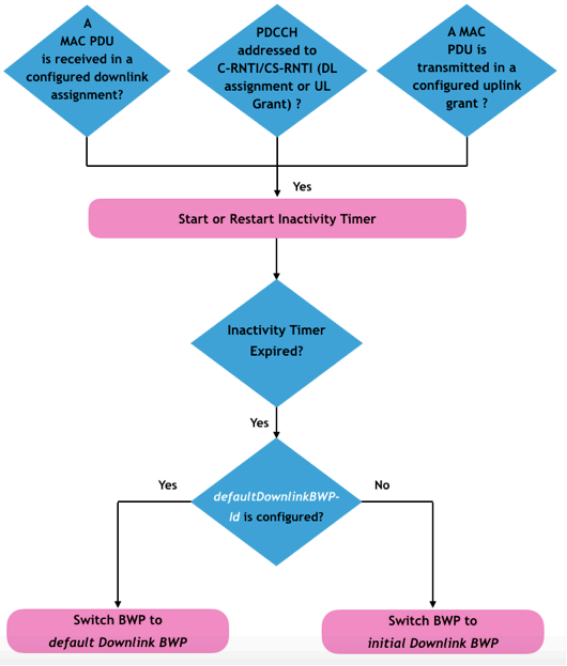


Figure 2.7: timer based BWP switching operation from [95]

have been introduced e.g. [98] [99]. In particular, [98], that is FlexRAN, and EmPower [99] represents the first SDN-enabled RAN architectures providing flexibility and programming in the radio access network. FlexRAN separates the data plane from the control plane of RAN according to the SDN logic and is implemented on top of the existing open-source LTE platform OpenAirInterface [100]. Instead, EmPower has been designed for WiFi and applied on srs LTE [101] to enable the concept of SD-RAN. As in [102] also, the SDN approach allows the network designers of SDN based radio resource management (SDN-RRM) to take over the distribution across the network of the resource radio manager radio resource management (RRM). This is because, the RRM is not able to sustain the optimal task of allocating resources, implementing handovers, managing interference, and balancing load between cells in a scenario resulting from the network densification in the path towards 5G networks [102]. It is worth noting that the main purpose of the SDN approach in the radio access network is the introduction of the flexibility and programmability of the radio segment itself. This flexibility can be introduced on different levels. This is because the element that we can consider to be programmable are those that have been modified along the path toward the 5G network. This means that, according to the SDN approach, we can assume that a SDN based controller monitoring the operational radio network condition thanks to its southbound APIs, can change nu-

merology, choose the right bandwidth part according to the equipped algorithms or even change the algorithms, to fulfil the required QoS. In this context, we proposed and present the following work.

2.4 Dynamic selection of numerology and bandwidth part for effective ORAN slicing

2.4.1 Introduction

The evolution process towards the 5G network, sees its final stage as a caterer of new services and improved technological assets. The new services 5G network is designed to support, are grouped in: Enhanced Mobile Broadband (**EMBB**), which leverages the high system capacity and the extremely expected throughput of 5G network for a better user experience, namely, augmented reality, streaming online gaming; Ultra-Reliable Low Latency Communication (**URLLC**), services leveraging reliable communications while fulfilling low latency requirements(self-driving car, industry automation); Machine to Machine Communication (**mMTC**), services characterized by a higher level of connection density, such as a smart city. For each category, 5G network has to guarantee specific target performance. For example, it is expected that for a URLLC service low latency and reliable low bit error rate are the main requirements to meet. It is expected that for an EMBB service the throughput is the main requirement to guarantee, whereas for mMTC service the high connection density should be taken into account as the first requirement. Additional requirements can be extracted from the coexistence of multiple services on the same network. All these requirements represent the starting point of the evolution process towards the 5G network.

A limitation of legacy mobile networks is the scarce flexibility, both software and hardware. This lack of flexibility can be observed in every network segment. For example, in radio access networks, the rigid and fixed structure of the radio frame according to the LTE system is not able to sustain all the expected scenarios. This is because this radio frame based on single numerology doesn't guarantee the required performance at all. In particular, for services with stringent latency requirements, the frame radio of LTE produces a Time Transmission Interval(TTI) that doesn't guarantee the target latency values of specific services. From here, the need for TTI reduction can be achieved with a flexible radio frame. From the software point of view, the scarce flexibility extends the time to in-field deployment innovation. The

rigid and monolithic architecture of legacy systems implies the development of such innovation in a distributed fashion. This approach is more time-consuming than a centralized one because the network is seen as multiple entities (nodes) which have to cooperate instead of a single-working entity. The improvement of network flexibility is the following step in the evolution process towards the 5G network. To provide flexibility, recent enabling mechanisms have been introduced.

Flexible Numerology (FN) is defined by 3GPP in [103]. Compared to LTE cellular system, the major difference is that the 5G radio supports multiple sub-carrier spacing. LTE has a fixed sub-carrier spacing of 15 kHz, whereas, 5G has the sub-carries space set $\{15, 30, 60, 120, 240\}$ kHz, corresponding to different values of numerology $\mu \in \{0, 1, 2, 3, 4\}$[1]. This allows us to define radio frame structure with shorter TTI as long as we keep selecting a higher value of numerology. By the way, this cannot be the only driving factor for the structure of the radio frame. In fact, according to the authors in [104], the selection of numerology should be taken considering for example processing delay of signalling activity related to the higher value of numerology and also traffic pattern characteristics.

Bandwidth parts (BWPs) represent another tool to increase the flexibility of 5G networks. Base stations are generally equipped with more computational resources compared to the UE. Due to this, signalling processing is more challenging in UE concerning a base station. Additionally, the FN introduces more complexity to signal to process. To address this complexity, the BWP has been introduced. Fundamentals on BWPs can be found in [105]. With the concept of bandwidth part, it is possible to divide the same carrier into multiple sub-carriers. Each of them has its numerology, thanks to FN, and other signal characteristics. The interested reader can find further details on [92]. This enables UEs to be configured to operate with the same carrier using customized numerologies and bandwidth sizes that fulfil user requirements in terms of latency, throughput, energy and spectrum efficiency.

2.4.1.1 Related Works

FN together with BWPS necessitates the study of switching mechanisms. In particular, the switching operation has been defined in 3GPP rel-15. The impact of the switching mechanism has been studied in [106]. In particular, the authors studied the impact of the BWP Inactivity Timer and BWP Switch Delay. They have shown that under certain conditions the reduction of frequency of switching operation is a key aspect for reduced energy consumption at the cost of higher latency and lower

[1] the sub-carrier spacing is defined as $\Delta f = 2^{\mu} \times 15$ kHz

throughput. In [107], the authors introduce an early form of bandwidth part adaption mechanism leveraging openair-interface and Flex-ran controller. Their results show that a relevant factor to be considered when defining an intelligent mechanism is the processing time wasted to evaluate the functionalities at the physical layer. This time is also linearly dependent on the number of UEs.

Instead in [108], the authors provide an iterative algorithm and a mathematical framework to address the bandwidth part adaption in a controlled manner. Also, this work represents an early mechanism that takes into account the coexistence of multiple services on the same network. In [109], the authors studied the impact of RAN slice bandwidth partitioning on slice performance but they do not focus on scheduling just selecting a simple Round Robin for their work.

Thus, the Bandwidth Part Manager (BWPM) is the software module responsible to handle the set of available bandwidth parts. It decides to place multiple BWPs on the same carrier to optimize the network's KPIs in different and dynamic manners.

The BWPM decides which BWP has to be bound to which network flow to match the requested QoS and it can properly modify the single BWP features, such as scheduling policies (Proportional Fair, Round Robin). An example of how alternative scheduling policies can introduce benefits (such as latency below 1 millisecond) can be found in [110]. Another example of an alternative scheduling policy can be also found in the work [111]. Here, the authors present a service discipline for BWP allocation based on the QoS requirements. In the work [112], the authors propose a dynamic BWP allocation scheme that switches between two multiplexing methods, dynamic multiplexing and orthogonal slicing, to minimize the impact of uRLLC traffic on eMBB traffic.

To increase flexibility, the rigidity of legacy system have to be addressed even from the software point of view. Recent enhancements have seen the adoption of the Software-Defined Network (SDN) approach to convert the software architecture of the Radio Access Network (RAN). In particular, the SDN-based approach converts the view of the RAN segment from a set of distributed entities to a single one. This implies new ways to design and deploy software applications which are, generally, easier concerning the past. In other terms, the RAN is more programmable and easier to control thanks to the single entity view which allows a centralized design and deployment of software applications.

A relevant application of the SDN approach in the evolution towards 5G network is provided in [113]. Here, the authors introduced the concept of a Software Defined Radio Controller to centralize the control logic of the legacy system. In particular,

it is expected that a pool of distributed controllers can introduce benefits in radio access network management. The authors assess that these controllers, according to the SDN approach, can communicate with each other thanks to East/West bounds while South bounds can monitor the RAN segment, which the controllers supervise, to provide a holistic vision of it. A further specialization of this controller has been introduced in [114]. The authors defined the "Radio Intelligent Controller (RIC)" to allow the adoption of machine learning mechanisms as well. In particular, this controller has been designed to handle two types of applications based on control-loop time. The RIC presents a "non-real-time" sub-module for applications requiring a control-loop time greater than 1 second. The "real-time" sub-module is used to run applications with control-loop time requirements below or equal to 1 second.

This architecture proposal is in line with the more recent effort of the O-RAN definition. The softwarization of RAN components imposes new technical challenges to avoid vendor-locking software chains. The O-RAN Alliance specification aims at enabling full open cooperation between vendors. Authors in [115] show a comprehensive overview of the O-RAN specifications, architectures and operations giving a glimpse of what is possible to do with O-RAN including xApps, rApps and AI/ML algorithms applied to O-RAN. In [116] authors shows how xApps and rApps may be used to provide a large-scale evaluation of ML-driven applications to control cellular networks in Colosseum and Arena test-beds.

2.4.1.2 Contribution

From the aforementioned research, the selection of the bandwidth part and the numerology have not been taken into consideration. Hence, to the best of our knowledge, we want to extend the study related to the process of bandwidth part selection, showing that a combination of the O-RAN approach and smart selection method allows us to achieve benefits in terms of data rate and latency for the end users. Hence, the contribution of this work is the following,

- we propose a software architecture of a radio controller acting as Bandwidth Part Manager (BWPM). This BWPM is equipped with two bandwidth part selection algorithms leveraging SDN functionalities to collect radio access network data. This data will be used to affect the bandwidth part selecting process to introduce benefits in terms of data rate and latency.

- we propose a bandwidth part and numerology selection algorithms that we named "Slice Aware" (SA) and compare it with a baseline "Physical Resources

Aware" (PRA) approach. The first algorithm adapts BWPs according to user requirements expressed in terms of latency, instead, the second selects the BWPs according to physical resource block assignment history.

- we perform a comparison of the two O-RAN Assisted BWPM strategies based on performance indicators such latency and experienced throughput.

2.4.1.3 Paper organization

The rest of the paper is organized as follows. In section **2.4.2.1 System Model**, we present the theoretical model of the key parameter indicators we studied, in sections **2.4.3.2 Slice Aware algorithm** and **2.4.3.1 Physical Resource Block Aware algorithm** we present the algorithms respectively. In section **Experimental Setup**, we present the experimental settings and describe our simulation campaign. In section **Result and Evaluation**, we argue the obtained results. In a section **Conclusion** we draw conclusions and the future directions of this work.

2.4.2 System Model

This section gives an overview of our system model architecture in 2.4.2.1. Performance metrics in 2.4.2.2 show our validation metrics.

2.4.2.1 System Architecture

Fig. 2.4.2.1 exhibits the system architecture under consideration. The architecture is composed of a set of gNBs which are controlled by an O-RAN service management and orchestration controller which is responsible for:

- monitoring users' performances
- monitoring bandwidth part and slices
- managing slices and bandwidth part life-cycles

In particular, Fig. 2.8 shows that, according to our *Control Loop*, the main functionalities are ***Data Collection and feedback*** and ***Policy and Control***. It is possible to implement an O-RAN-assisted BWPM xApp that, following the *Control Loop* and according to the network's condition, can alter the configuration of already available BWPs or possibly add new ones for a specific service. This way it is possible to extend the network slice up to the end user, through the radio segment. This allows us to overcome the issue of QoS's worsening related to fixed bandwidth part

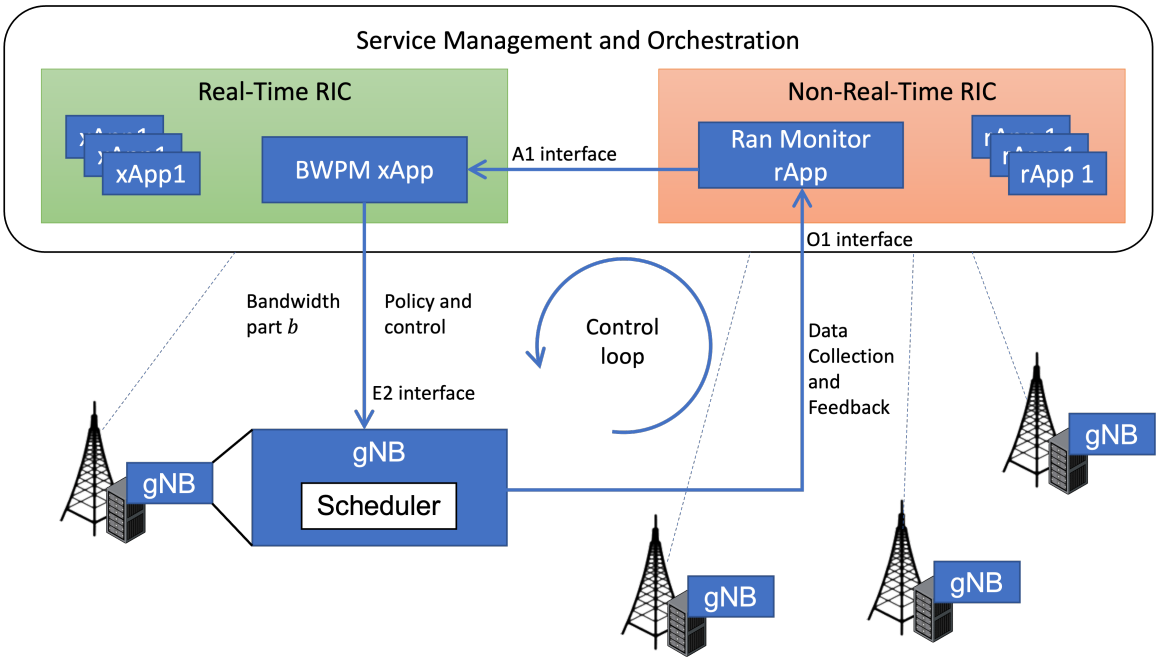


Figure 2.8: System Model Architecture

management which may not be suitable for supporting 5G-related network services because it doesn't consider the changes over time of the network's operating condition. In detail, the ***Data Collection and feedback*** is executed by the customized **Ran Monitor xApp** software component collecting state S by the gNBs that it supervises through the O1 interface. The observed state S represents the network operating condition of RAN and it is defined as follows: let $U = \{1, 2, ..., N\}$ be the set of N users and B the set of available BWPs in the base station of a cellular system. Then, S is defined as $\{B, (L(B), P(B))_{i \in U}, T(B), H(B)\}$ where:

- $B = \{b_j | j \in \{0, 1, 2, 3, 4\}\}$is the set of available bandwidth parts with progressive numerology $\mu \in \{0, 1, 2, 3, 4\}$. For B we define total ordering property: $\forall (b_k, b_j) \in B^2 \wedge b_k \neq b_j,\ b_k < b_j \iff \mu(b_k) < \mu(b_j)$. Also, as interchangeable notation, $\mu(b_j) = \mu(j)$ represents the numerology of bandwidth part $b_j \in B$ and $b_j = j$ represents the bandwidth part $b_j \in B$;
- L(B) is the measured L latency $\forall b_j \in B$;
- P(B) is the set of transmitted packet $\forall b_j \in B$;
- T(B) is the set of threshold value $\forall b_j \in B$;
- H(B) is the history of assigned physical resource block $\forall b_j \in B$;

- $(L(B), P(B))_{i \in U}$ is the registered L_{UP} latency and transmitted packet for each $j \in B$ and for each user in $i \in U$.

Such S represents the holistic RAN state that is monitored by the **RAN Monitor rApp** of our prototype O-RAN assisted BWPM. Then, S is converted in a special format (S_{SA}, S_{PRA}) and provided to customized **BWPM**. The customized BWPM, instead, executes the ***Policy and Control*** enforcement. With it, the BWPM gets the S_{SA} or S_{PRA} state and performs a BWP selection algorithm. The BWP selection algorithm produces as output b[2], which is the BWP that the scheduler has to use for the transmission of the current packet. Further details on S_{PRA} will be provided in Sec. 2.4.3.1, instead, for the S_{SA} we refer to Sec. 2.4.3.2.

2.4.2.2 Performance Metrics

In this work, we focus on the study of downlink performance. Packets are first processed at the Core Network and then sent to the gNBs for transmission to the mobile users. Since we are interested in the evaluation of the impact of different BWPM strategies on the RAN performance we assume a zero contribution for core components [117]. This is reasonable because it is expected that the next RIC should be located near the gNBs, increasing network efficiency [118], and every decision-making process applied to RAN has to be based on RAN KPIs such as latency radio or more specific ones.

Fig. 2.9 provides a graphical representation of the adopted latency modelling. The gNB is equipped with a queue where packets are stored before being transmitted. When a packet has to be transmitted to scheduled users, it is converted before in a transport block or possibly a set of code blocks[3] and then is sent to the specific users.

The packet flow of the N UE packets coming from Core and stored in the queue is modelled according to a traffic model based on what we assume to be Poissonian. When a packet is removed from the queue, it has to be processed by the gNB. The radio protocol stack impacts the radio latency, in a sense that, the overhead of the internal processing procedure that allows a packet to be converted to a transport block or possibly in a set of code blocks, introduces delays that have to be considered to properly evaluate the radio latency. Then, after the packet has arrived at the user side, it has to be processed from UE. Considering the aforementioned delays

[2] $b = SA(S_{SA})$ or $b = PRA(S_{PRA})$

[3] This depends on the size of the packets that have to be transmitted

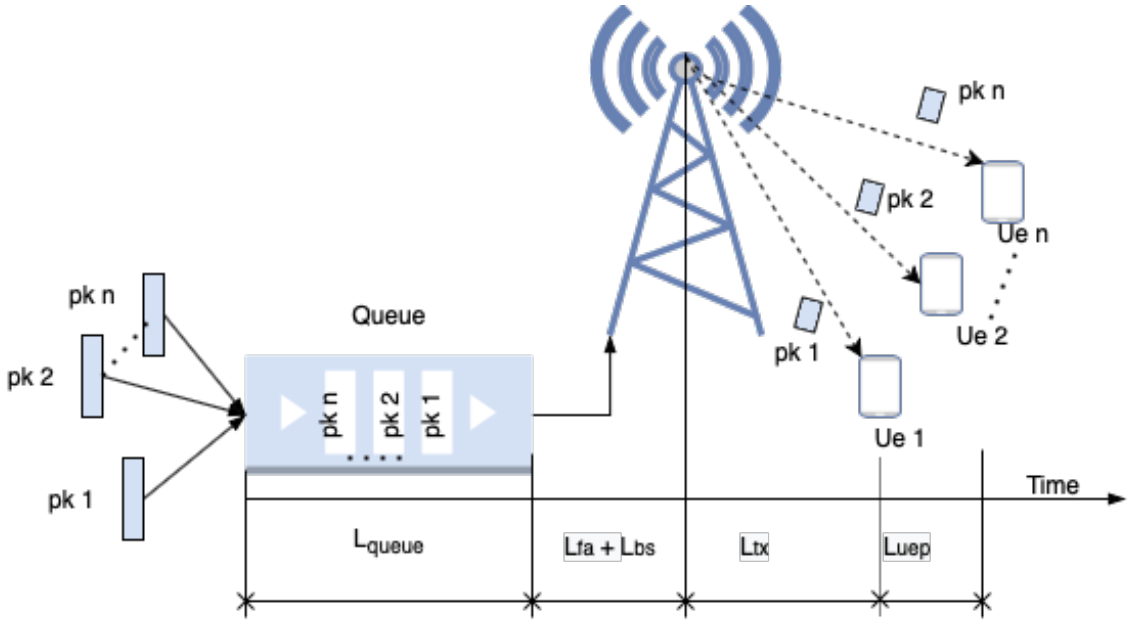


Figure 2.9: Latency model

contributing to the radio latency, according to [117] also, we can model such latency as follows:

$$L = L_{queue} + L_{fa} + L_{tx} + L_p^{bs} + L_p^{ue} \tag{2.2}$$

where:

- L_{queue} is the time spent by a packet in a buffer/queue; it depends on the number of users that can concurrently transmit.
- L_{fa} is the delay due to the frame alignment which depends on the frame structure and the duplexing mode (FDD, TDD).
- L_{tx} is the transmission delay based on the radio channel condition, payload size, available resources, transmission errors and re-transmissions;
- L_p^{bs} and L_p^{ue} are the packets' processing delays at gNB and UEs respectively; these depend only on the machine's capabilities.

and according to [119], corresponds to the one-way transit time between a packet being available at the IP layer of the sender UE/BaseStation and the same packet being available at the receiver BaseStation/UE. L depends on several factors, and the goal of minimizing it can be achieved employing different techniques [117] belonging

to different areas such as advanced multiple access schemes, advanced modulation, and coding scheme, packet/frame structure, etc.

As stated in Sec. 2.4.1, in this work we also considered the experienced throughput of each user. This key parameter indicator has been defined as follows: let's consider the user's flow i, the set B of the available bandwidth parts indexed with j, the number of transmitted packets with bandwidth part $j \in B$ for the i^{th}-flow $P(j)_i$. Then, the total number of transmitted packets T_i of the flow i is given by the sum over B set during the simulation as follows:

$$T_i = \sum_{j \in B} P(j)_i \tag{2.3}$$

. Then, for each of the i^{th}-flow's packet k, we consider also the partial delay δ_k. The sum of partial delays Δ_{T_i} can be defined as follow:

$$\Delta_{T_i} = \sum_{k=1}^{T_i} \delta_k \tag{2.4}$$

The value of the experienced throughput T_h^i for the user i is defined by the following:

$$T_h^i = \frac{T_i}{\Delta_{T_i}} \tag{2.5}$$

2.4.3 Bandwidth-part selection algorithms

In this section, we describe the two algorithms we developed as core decision-making functions for our O-RAN-assisted bandwidth part manager. Both algorithms take as input the relevant information state we collected from the observed radio access network. As output, they provide the bandwidth part that the scheduler must use to send the current packet to the specific user. Further details are provided in the dedicated section for each algorithm.

2.4.3.1 Physical Resources Aware algorithm

This method considers the Physical Resource Blocks (PRBs) assigned to transport blocks (TBs). The idea behind this strategy is to always prioritize BWPs with higher numerology since they are expected to introduce lower latency (due to higher subcarrier spacing and lower symbol time). The packets are assigned to lower numerology BWPs only when the higher ones are saturated and no resources are available.

The PRB is the smallest radio resource which can be assigned to transmit data. It comprises 12 sub-carriers spanning in one slot with different slot lengths according

to FN. Instead, the TB consists of a set of bits with a specific size. In particular, the medium access control (MAC) layer organizes the data into the TB and transmits it to the physical layer(PHY) using PRBs to transmit it to the receiver. Additionally, in 5G NR, the maximum transport block size is 1,277,992 bits [120]. When the TB's size exceeds the thresholds, a TB is segmented in multiple code blocks (CBs) as shown in Fig. 2.10. Further details on TB processing are described in [121]. The code block consists of up to 8448 bits. Both the transport block and the code block have a cyclic redundancy check (CRC) attached. Due to the difference in the size of the TBs and CBs, the CRC processing scheme for TB is slightly different for CBs. Both TB and CB need PRBs to be transmitted to the receiver. We define S_{PRA} the state provided to the PRA algorithm.

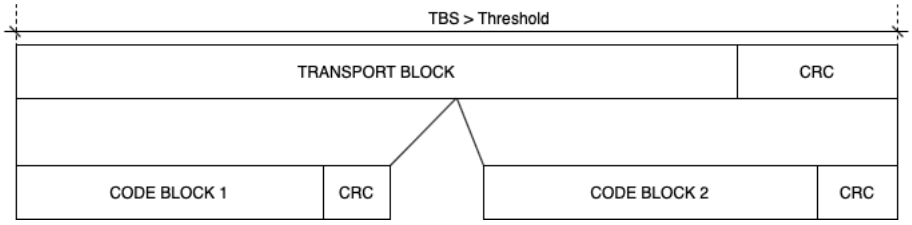


Figure 2.10: Segmented Transport Block

The S_{PRA} state's components that PRA analyses are $\{T(B), H(B), B\}$. Let's define $MAX_{\rho(j)}$ as the maximum number of physical resource blocks which can be assigned given specific numerology associated with the BWP j and ratio $r \in (0, 1]$. For each $j \in B$ we define threshold value $t(j) = r \cdot MAX_{\rho(j)}$ used to limit the usage of BWPs and the $T(B) = \{t(j)|j \in B\}$ the threshold values set in accordance. Given $\boldsymbol{T_w} = 1, 2, 3, ..., t_w$, let's $H(B) = \{h(j)|j \in B\}$ the set where $h(j) = \{\rho_k|k \in \boldsymbol{T_w}\}$ is the PRBs assignment-based history indicating ρ_k as the amount of assigned PRBs to the current k TB or possibly CB. At this point, for each $j \in B$, we evaluate the bandwidth part occupation B^j_{Occ} as:

$$B^j_{Occ} = \frac{\sum_{\rho_k \in h(j)} W_{\rho_k} \cdot \rho_k}{\sum_{\rho_k \in h(j)} W_{\rho_k}} \tag{2.6}$$

The bandwidth part pick is based on B^j_{Occ}, in particular, the method checks whether this value is less than the corresponding threshold value $t(j) \in T(B)$. This method keeps selecting the same BWP, if and only if, $B^j_{Occ} < t(j)$ holds, otherwise it will select the first next eligible BWP. By the way, this method will rarely change the BWP

Table 2.1: Notations for PRA algorithm

Notation	Description
S_{PRA}	Observed state PRA is interested in.
B	BWPs set.
$T(B)_i$	Set of Threshold values.
$t(j)$	Threshold value for $j \in B$.
r	Ratio value $\in (0, 1)$.
$MAX_{\rho(j)}$	Maximum amount of assignable Physical Resource Blocks specific for $j \in B$.
$H(B)$	History of PRBs assignements for all BWPs.
$h(j)$	History of PRBs assignments for $j \in B$.
$\boldsymbol{T_w}$	Time window.
B^j_{Occ}	Bandwidth part occupation based on $H(B)$.
W_ρ	Weight of specific assignment $\rho \in h(j)$.
ρ	Amount of assigned PRBs to the current TB or eventually CBs.
$\rho_{k \in \boldsymbol{T_w}}$	k^{th}-amount of assigned PRBs to k^{th}-TB or possibly k^{th}-CBs.
bwp	Selected BWP $j \in B$.

because the assigned PRBs to TBs decrease as long as the system keeps transmitting. The system is consuming data during transmission and this implies smaller and smaller TBS corresponding to the smaller amount of PRBs that have to be assigned.

To address this situation and force the method to change BWP, we introduce a time window $\boldsymbol{T_w}$ and the ratio r and we set different weights W_ρ for the PRBs assignments. In particular for $\rho = MAX_{\rho(j)}$ the $W_\rho \gg W_{\rho'}$ with $\rho' \neq MAX_{\rho(j)}$. It is clear that the change of BWP, at this point, stems from the number of full-available PRB assignments, namely $\rho = MAX_{\rho(j)}$, we have in $h(j)$ and the settings of $\boldsymbol{T_w}$ and ratio r.

It is worth noting that at the initial stage, namely when we have $|h(j)| < t_w$, the PRA keeps using the first configured and available BWP j, that is the $bwp = 1$ with $\mu = 0$. Also in a full swing condition, if $B^j_{Occ} < t(j)$ doesn't hold for all available BWPs j, the system keeps using the same used BWP in the last transmission. This information corresponds to the $j \in B$ such that $\rho_{t_w} \in h(j)$.

2.4.3.2 Slice Aware algorithm

We call S_{SA} the part of the holistic RAN state S the Slice Aware algorithm is interested in. It is defined as: $S_{SA} = \{(L(B), P(B))_{i \in U}, L, B, \}$ where:

- $L(B) = \{L(j) | j \in B\}$ represents the average of registered L_{UP} for each BWP j;

Algorithm 1: Physical Resources Aware

Result: *bwp*
$S_{PRA} = \{(T(B), H(B), B\}$;
$\boldsymbol{T_w}$;
$bwp = 1$;
$Find = false$;
for *each* $j \in B$ **do**
  evaluate B^j_{Occ} according to (2.6);
  if $B^j_{Occ} < t(j)$ **then**
    $bwp = j$;
    $Find = true$;
    break;
  end
end
if $Find = true$ **then**
  return *bwp*;
else
  return $bwp = j \in B$ s.t. $\rho_{t_w} \in h(j)$;
end

- $P(B) = \{P(j)|j \in B\}$ represents the number of transmitted packets for each BWP j;
- $(L(B), P(B))_{i \in U}$ represents the user-based data composition of the $L(B), P(B)$ for each user i;
- $L = \{L_i|i \in U\}$ represents the set of user-based L_{UP} requirements.
- B represents the set of available and configured BWPs;

The idea behind SA is to find the first bandwidth part with the highest L_{UP} satisfying the user's latency requirement $L_i \in L$. In other terms, given full swing condition of the cellular system and the set BWP, for each out-going packet belonging to the data flow of users i, the method selects BWP j such that $L(j)_i < L_i$ and $L(j)_i = max\{L(B)_i\}$[4].

The full-swing condition is determined by the guarding expression of the first if-statement acting as the filter. Basically, given a user i, it checks if the number of

[4]To address the situation where you have multiple and equal values for the max values in $L(B)$ the method resolves it with the pick of the BWP with lower numerology. This especially happens at the initial stage where the registered values of L_{UP} are equal to 0 hence suitable candidates to be max values of $L(B)$

Algorithm 2: Slice Aware (SA)

Result: $j \in B$
$S_{SA} = \{(L(B), P(B))_{i \in U}, L, B, \};$
$\overline{X} = \overline{1};$
$bwp = 1;$
$\hat{P} = C \cdot |B|;$
if $P(\overline{X}[i])_i > \lfloor \hat{P}/(|B| - \overline{X}[i]) \rfloor$ **then**
  $j = argmax_j \{L(j)_i \in L(B)_i | L(j)_i \leq L_i\};$
  if $j \in BWP$ **then**
    $\overline{X}[i] = j;$
  else
    $j = \overline{X}[i];$
  end
else
  $j = \overline{X}[i];$
end
return j;

transmitted packets with BWP $\overline{X}[i]$ is greater than $\lfloor \hat{P}/(|B| - \overline{X}[i]) \rfloor$ where $\hat{P} = C \cdot |B|$ with $C \in \mathbb{N}$. This condition depends on the available BWPs in the system and $\hat{P}$. The definition of $\hat{P}$ comes up from the analysis of the cellular system's behaviour. The result of this analysis is the existence of different transient times for each BWP. In particular, every time we change BWP, the variance of the registered L_{UP} is way higher than the one in a full-swing condition of the system. This means that to properly choose the next BWP, we need to avoid transient time. To account this situation we estimates a filter value corresponding to $\lfloor \hat{P}/(|B| - \overline{X}[i]) \rfloor$.

This guarantees SA selects BWP based on filtered registered L_{UP} values. If $P(\overline{X}[i])_i > \lfloor \hat{P}/(|B| - \overline{X}[i]) \rfloor$ holds then SA starts searching a new BWP or eventually keeps selecting the old one, otherwise the selecting process of the BWP is immediately delayed after the transient time. To support the selection process we additionally define the vector $\overline{X}$ containing the used BWP in the last transmission. In particular, this vector has $|\overline{X}| = |U|$ and $\overline{X}[i] \in B$ represents the used BWP in the immediately previous transmission related to the traffic flow of user $i \in U$.

Additionally, if the evaluation process of Slice Aware's S_{SA} takes into account the bandwidth part with the highest numerology, namely 5^{th} BWP[5], the immediate result is the 5^{th} BWP. This means that the system either for the first time selects as new BWP the one with the highest numerology or keeps using the same of the last

[5]It is the BWP with $\mu(5) = 4$

Table 2.2: Notation for SA algorithm

Notation	Description
S_{SA}	Observed state SA is interested in.
B	Set of available BWPs in the system.
L	Set of L_{UP} requirements.
L_i	$i \in U$-based L_{UP} requirement.
$L(B)$	B-based registered L_{UP} latency values.
$P(B)$	B-based number of transmitted packets .
$(L(B), P(B))_i$	$i \in U$-based registered L_{UP} latency and B-based transmitted packets
$(L(j), P(j))_i$	Registered L_{UP} latency and transmitted packets with $j \in B$ for i^{th}-user
$\overline{X}$	Vector of used BWPs $j \in B$ in last transmission.
$\overline{X}[i]$	$j \in B$ used BWP in last transmission for user $i \in U$.

transmission which was the 5^{th} BWP. In both cases, SA selects the 5^{th} BWP which is the one which is expected to be characterized by lower latency. This cannot be improved simply because BWP with higher numerology than 4 has not been defined [103] and lower BWPs are not enough to satisfy the condition $L(j)_i < L_i$. It is worth noting that if the condition $L(j)_i < L_i$ doesn't hold for the 5^{th} BWP, SA will keep using the BWP with the highest numerology that has been stored within the vector $\overline{X}$ the first time it has been used. In Table 2.2 are summarized all the used parameters.

2.4.4 Results and Evaluation

2.4.4.1 Simulation Scenario

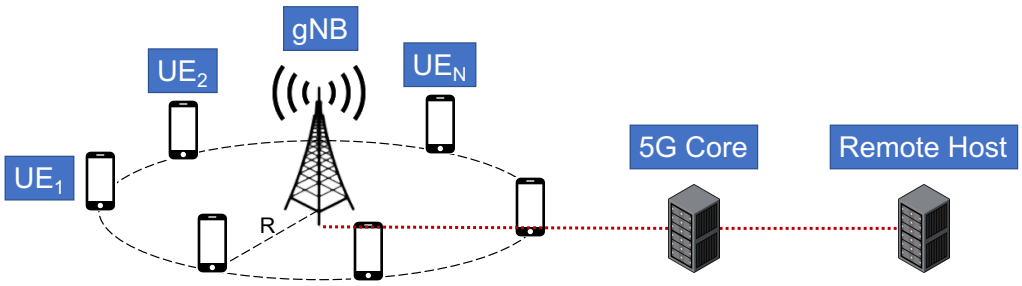


Figure 2.11: System Cellular

We consider the simulation scenario represented in Fig. 2.11. According to the service area dimension defined in [122] and [123], and urban macro scenario characteristics, we consider a single Base Station (BS) with users being randomly placed

Table 2.3: Physical Settings

Physical Parameters	Values
BS Height	25 m.
Channel condition	Line of sight.
Scenario	Urban macro.
Shadowing	no.
Ue height	1,5 m.
BWPs	5.
Numerology	{0,1,2,3,4}.
Central frequency	28 GHz.
Bandwidth	100 Mhz.
Scheduler	Proportional Fairness.

around it at a distance R. Also, according to [123], the height of the base station is 25 meters, instead, to simulate a hand-managed device the height of user equipment is 1,5 meters. We assume the absence of shadowing effects for the propagation model and the Line of Sight channel condition for the path loss model. The base station is equipped with a component carrier and 5 full-featured BWPs with progressive numerology according to the definition of BWP set in Sec. 2.4.2.1. The default BWP is the 1$^{\text{st}}$ with $\mu(1) = 0$, instead, the others have higher numerology values and can be used to transmit additional traffics and benefit from them [92]. The main features of BWPs include a central frequency of 28 GHz, a bandwidth of 100 MHz and Proportional Fairness as a scheduling policy. UEs are configured such that they can process the signals coming from the base station. So, based on network operational condition needs, they can process all the BWPs. The table 2.3 lists all the physical settings adopted. We consider 3 sets of users U_1, U_2, U_3 such that $|U_1| = 10$, $|U_2| = 20$,$|U_3| = 30$. Also, we consider users belonging to three slices $S_l = \{a, b, c\}$ differing among themselves for the L_{UP} requirements. Slice **a** imposes L_{UP} requirement to be maximum 2ms, slice **b** sets L_{UP} requirement to 3ms, and slice **c** has L_{UP} requirement equal to 4ms. This requirement doesn't change over time. Then each pool of users is uniformly distributed in the classes. The traffic template [122] corresponds to the URLLC discrete automation. This type of traffic template defines a packet size D which is smaller than 1358 bytes and greater than 256 bytes. Therefore we set $D = 1000$ bytes. Further details can be found in [122].

Based on Sec. 2.4.3.1, we need to set the time window and the ratio r to affect the usage of BWPs. To simulate stress conditions and force the algorithm to change the BWP, we set $\boldsymbol{T_w} = 1$ and $r = 99\%$. In this way, the bandwidth part occupation

is only related to the last physical resource block assignment. Thus, the likelihood of BWP change is higher than the situation with larger $\boldsymbol{T_w}$. This because is very likely $B_{Occ} = 100\%$. Based on Sec. 2.4.3.2, given the set of L_{UP} and established that they don't change over time, we need to define the value of our packet filter, that is the value of constant C. In this work, $C = 10$.[6]

The simulations are organized as follows: for each pool of user U_1, U_2, U_3 we executed the first simulation with the algorithm PRA as the BWPM selection method. Then, we executed a second simulation with SA as the BWPM selection method. The simulation time is 5 minutes for all six experiments. All the simulations were executed on Ns3-simulator which was integrated with a millimetre wave radio environment [104]. Table.2.4 lists all the simulation settings.

2.4.4.2 Results

In this section, we present and discuss the obtained results. The obtained data from two simulations are grouped and analyzed by L_{UP} requirements. Figs. 2.12 to 2.14

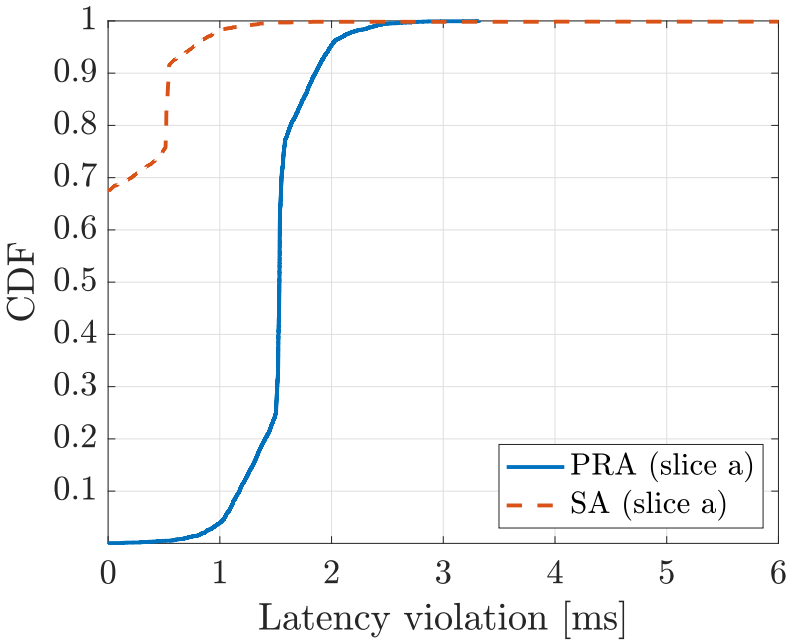


Figure 2.12: Slice a latency violation

analyze the experienced latency violation for different classes of users and compare the performance of the 2 proposed BWPM strategies, i.e., SA and PRA. The latency

[6]This value has been obtained experimentally during the evaluation of system behaviour's transient.

Table 2.4: Simulation Settings

$\boldsymbol{U}$	$\lvert\boldsymbol{U}\rvert$	
U_1	10	
U_2	20	
U_3	30	
User Class	$\boldsymbol{L_{UP}}$	**Note**
a	2 ms	The users are uniformly distributed on all classes
b	3 ms	
c	4 ms	
TFT	**D**	**Note**
Discrete Automation	1000 B	$256B \leq D \leq 1358B$ [122]
Algorithms	**Settings**	**Note**
PRA	$\boldsymbol{T_w} = 1$, $r = 99\%$	Simulation time 5 minutes
SA	$C = 10$	

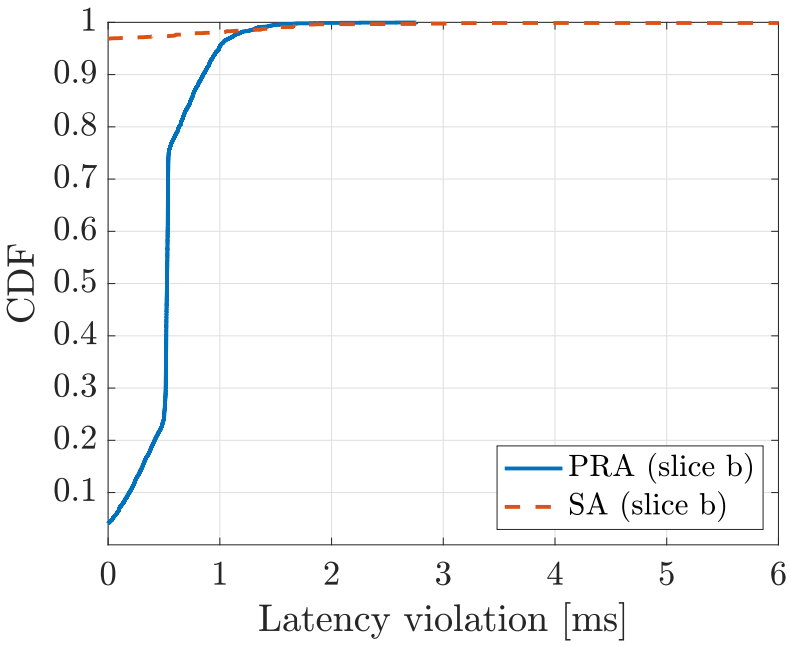


Figure 2.13: Slice b latency violation

violation is defined as $L_v = |L_{UP} - Li|$ where L_{UP} is the registered user plane latency, while Li is the user plane latency requirement for the user i. It is worth noting that the probability value corresponding to a latency violation equal to zero, that is the absence of latency violation, is 0.7 with SA (dashed line) and 0 with PRA for the slice a user. For slice b the values are almost 1 for SA (dashed line) and less than 0.1 for PRA. Analogous results can be observed for slice c users. As expected, PRA represents a feasible solution, but our O-RAN-assisted BWPM equipped with SA as the BWP selection method performs better than the PRA method.

Figs. 2.15 and 2.16 we show the impact of the number of users on latency and latency violation for the Slice Aware strategy. In Fig. 2.15, we plot the experienced latency based on the number of users. As one can see, the higher the number of users, the higher the probability of having high latency values. Additionally, the indented curve trends are related to the user distribution onto the set C_l. This is what we expect because the network load corresponding to 30 users is higher than the network load of 10 users. In other words, the amount of data that the base station has to process grows with the number of users and this means that the queuing time discussed in Sec. 2.4.2.2 grows as well. Also, this can be observed from the latency violation viewpoint, as in Fig. 2.16 where we plot latency violation for the different number of users. This plot shows that the probability of having zero violation during

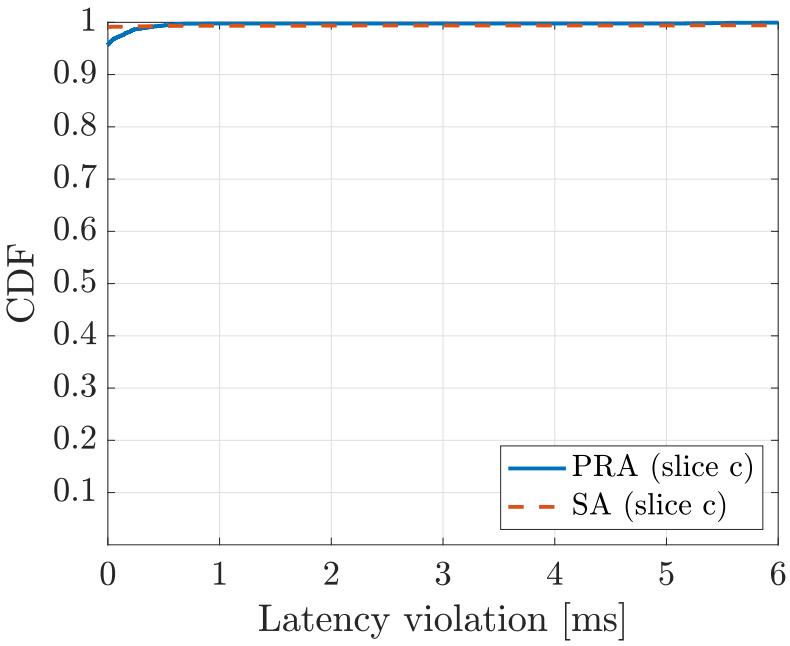


Figure 2.14: Slice c latency violation

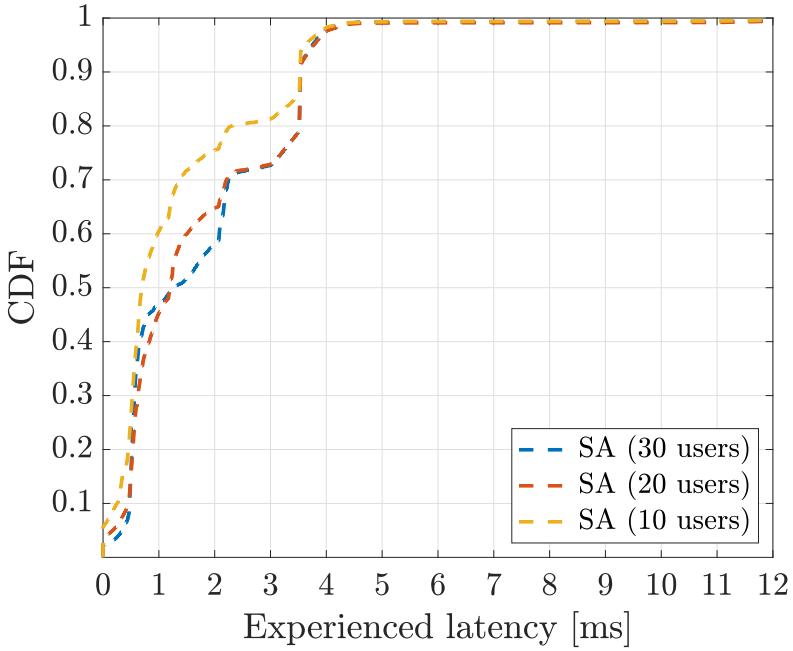


Figure 2.15: Experienced latency vs. Number of users

transmission is above 0.5 and decreases with the increase of users. Fig. 2.17 shows the comparison between SA and baseline method PRA by only considering U_3, i.e. 30 users scenario. In Fig. 2.17, we compare the obtained latency values with the

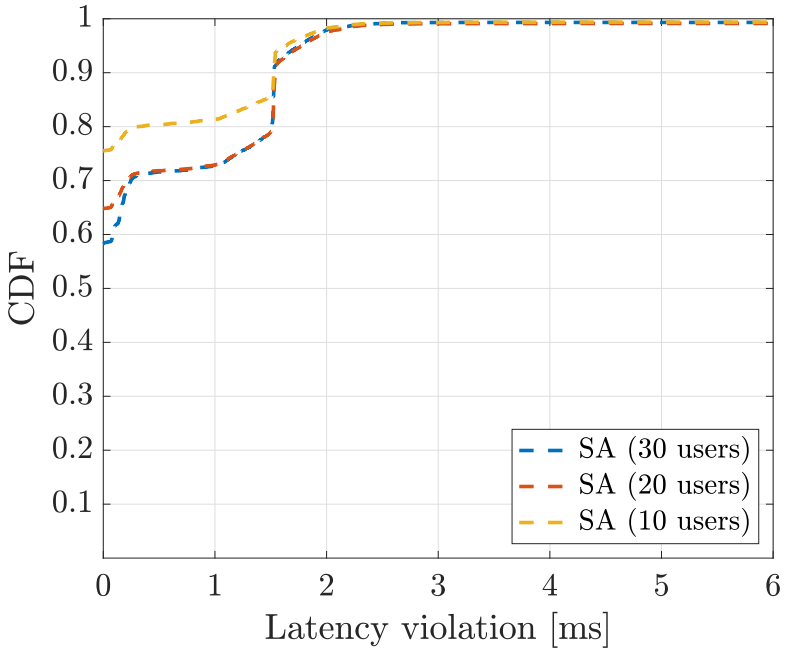


Figure 2.16: SA latency violation user pool-based comparison

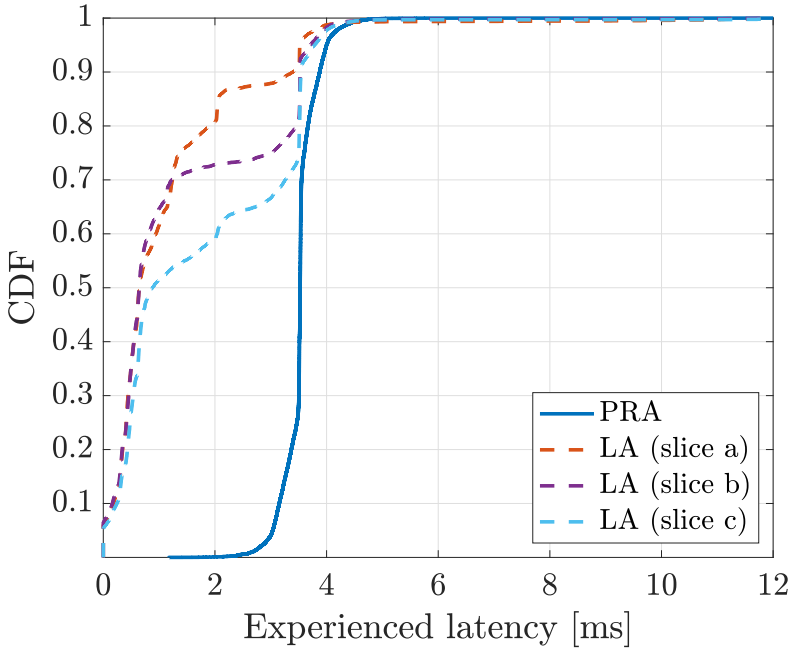


Figure 2.17: Experienced latency SA and PRA comparison

PRA method, and the obtained latency values with SA, represented with dashed lines and related to all slices. Results show that, although the PRA strategy tends to utilize higher numerologies which are characterized by lower latency, the SA strategy

performs better in terms of experienced latency. This is because it does not offer low latency BWPs to users with higher latency constraints thus, efficiently distributing traffic among BWPs and avoiding queuing of traffic at the gNB.

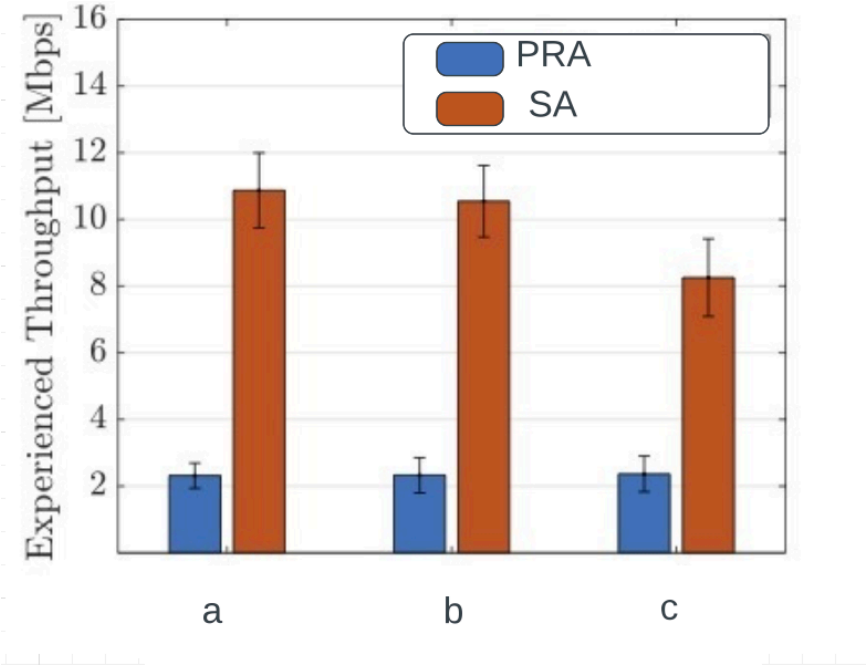


Figure 2.18: Throughput comparison between PRA and SA

According to the definition of users experienced data rate of Eq.2.5, we can derive the following consideration, that is, for the same T_i, Δ_{T_i} is the main responsible for the variation of T_h^i. A flow of packets transmitted with BWP with high numerology values produces δ_k very small respect to one with BWP with lower numerology. This produces Δ_{T_i} values in accordance. So, for the same T_i, we have T_h^i very high for SA concerning PRA, simply because the BWPM equipped with SA tends to use BWP with high numerology to fulfil the user L_{UP} requirements. The consequence of this is that the registered throughput with SA, on average, is 5 times higher than the registered one with PRA, as shown in Fig. 2.18.

2.4.5 Conclusion

The 5G network is continuously evolving. It is expected that this flexible network environment will include different technological solutions to provide any kind of benefit. Recent enhancements in 5G network will leverage enabling mechanisms, such as flexible numerology and bandwidth parts. Additionally, it is worth mentioning Software Defined Networking approach together with virtualization and O-RAN will

be beneficial as well. In this work, we investigated the dynamic selection of the bandwidth part. In particular, we propose an O-RAN-assisted BWPM equipped with BWP selection algorithms called Slice Aware and Physical Resources Aware xApp. The obtained results show that the BWPM equipped with SA xApp performs better than PRA xApp in terms of experienced latency although the latter tends to prioritize higher numerology with lower latency. The future direction of this work will take into account an alternative version of O-RAN-assisted BWPM, according to the newest enhancements of Radio intelligent controllers including machine learning approaches such as Reinforcement learning at the edge as rApp. Moreover, it is also worth investigating the performance of the considered system in the case of multiple base station scenario and the how they are affected if we considered different height based-positioning of the base stations.

2.5 Summary

In this chapter, we introduced the main radio access technologies allowing the evolution of the radio access networks towards 5G. We are referring to the flexible numerology, bandwidth part and its management. The review of the state of the art allowed the definition of the problem that has been addressed in the work titled **Dynamic selection of numerology and bandwidth part for effective ORAN slicing**. By experimental activity, we showed that dynamic selecting mechanisms for numerology and bandwidth parts, supported by the SDN networking paradigm, introduced tangible benefits which can meet the 5G QoS requirements for the considered services. This is in contrast with the results obtained by their static alternatives.

Chapter 3

Resource Management in a Cell-Free Massive MIMO scenario

In this chapter, we introduce Resource Management in a Cell-Free Massive MIMO scenario. Due to its huge performance capabilities, the Cell-Free Massive MIMO is the most promising radio access network configuration that can be used to fulfil the requirements stemming from the network densification together with the new 5G services. This kind of configuration has specific resources that have to be efficiently managed to fulfil the expected quality of services of new 5G scenarios including the coexistence of multiple services. Due to this consideration, this chapter has been arranged as follows: in the **sec: 3.1 Introduction and Motivation**, we introduce the cell-free massive MiMo and we explain the motivation leading to the deployment of this configuration. In the **sec: 3.2 User-Centric Cf-MiMo**, we present a detailed description of cell-free massive MiMo configuration and it has been arranged as follows:

- in **sec: 3.2.1 System model for Ul and Dl**, we present the most used theoretical fundamentals modelling the Ul and Dl transmitted signals;
- in **sec: 3.2.2 Resource Allocation: Pilot Assignment Algorithms and Power Control/Allocation Schemes**, we review the most used algorithms and schemes to allocate the cell-free massive MIMO resources, that are pilot sequences and power.

In **sec:3.3 Multiplexing URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity**, we present the current progress of the work named "Multiplexing URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity". Here, we define a theoretical framework for performance evaluation of such a

system, based on closed form expressions for the spectral efficiency and the throughput of eMBB and URLLC services in this new asset. We show the advantages of the cell-free massive MIMO to support such heterogeneous services while guaranteeing their specific QoS requirements. Here, we propose a way to handle the resource management of the cell-free massive MIMO to sustain the coexistence of multiple services.

3.1 Introduction and Motivation

The mobile networks have been designed for providing devices access to a variety of data services anywhere in a wide geographical area. The main provided service in this area was voice calls. However, this service has been taken over by packet transmission. So the service quality of the network is represented by the level of the data rate that can be delivered in each location within the coverage area. The propagation environment determines the coverage area. Since the received signal power decreases quadratically, with the propagation distance, a traditional mobile network infrastructure consists of a set of geographically distributed APs that the UEs can choose between. The APs are deployed at elevated locations to improve the coverage area. The cellular network is composed of APs and UEs that select the strongest one to connect to the network. The location of the UE for which the AP is selected is called a cell. In this architecture, the main reason for the performance degradation experienced by the users is that the user close to AP experiences better SNR with respect to the ones close to the edge of the cell and also it will experience interference related to the neighbourhood. Hence the SINR is lower than SNR. Since the data rate is an increasing function of the SNR, in each cell each user could experience large data rate variations. The best conditions are near the centre of the cell and the worst are at the edge. The legacy cellular network was designed with the purpose of avoiding call drops. These call drops could be avoided by providing SNR above a specific threshold in each cell. Anyway, there weren't benefits to being far from the thresholds. When the cellular network started to sustain the data transmission the main property changed. In particular, the new property is that the UEs request the same data services everywhere in the coverage area. Hence, the UEs near the centre of the coverage only need to be connected part of the time, while the UEs at the edge of it must be turned on for a much larger fraction of time. So, at a given time instance, the majority of active UEs are at the cell edges and their performance will determine how the customers perceive the QoS of the network as a whole. To sustain

this property, massive MiMo technologies have been deployed. With these technologies, the cellular networks started benefiting substantially data rate improvements. Despite these gains, in the cells remained data rate variations not negligible. These variations, in a society where it is supposed to have a ubiquitous wireless connection, are not admissible. Moreover, Since the legacy cellular network was designed for low-rate voice call drop service instead of uniform high data rate service, they are not eligible for the new generation of services of 5Gns. This uniform high data rate everywhere pushed the researcher to consider new cellular architectures. The promising architecture, in this context, is represented by the Cf-maMiMo.

3.2 User Centric CF-MiMo

As described in [81], **the Cf-maMiMo** is defined as *ultra dense network where the APs are cooperating to serve the UEs by joint coherent transmission and reception while making use of the physical layer concepts from the cellular Massive MIMO area.* The fig.3.1 represent this MIMO system. As you can see from fig.3.1, the Cf-maMiMo

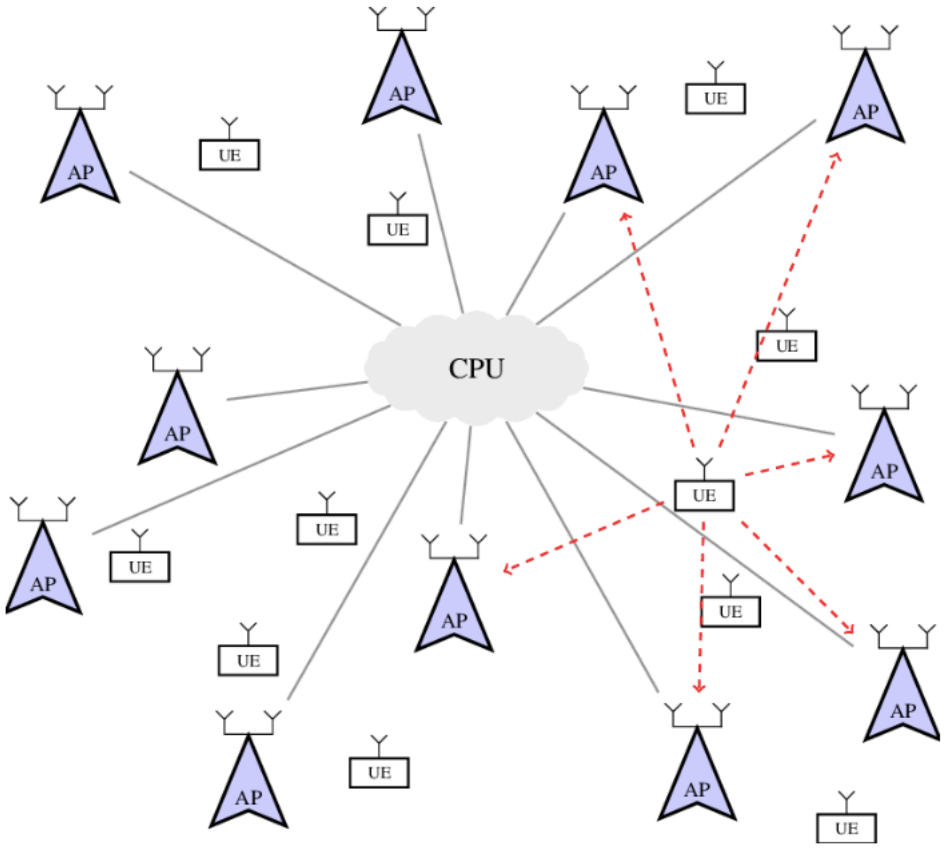


Figure 3.1: Cf-maMiMo ideal concept from [124]

has L APs, each equipped with N antennas, that are geographically distributed over the coverage area. Also let $M = N \times L$ denote the total number of AP antennas in the network. The APs are jointly serving K single-antenna UEs. More precisely, each UE is communicating with a subset of the APs, which is selected based on the UE's needs. The APs are connected via fronthaul links to CPUs, which facilitate the AP coordination. The mathematical definition of an ultra-dense network is $L \gg K$ implying also $M \gg K$, that is the total number of AP antennas is much larger than the total number of UEs as in a conventional massive MIMO system; The main operating condition in the cell-free network is that whenever the APs focus the transmission towards a particular UE, the focus area will be sufficiently small so that no other UE is receiving experiences high interference. The cooperation of APs refers to User-Centric features of the Cf-maMiMo systems. Given the main idea of all the APs serving the specific UE, in practical systems, it is reasonable to think that the UE is served by a subset of APs that is called a cluster. This is the idea of allowing the **user based APs cluster composition**. For different users, we could have different clusters that possibly could overlap. This overlapping afterwards is because cluster membership of specific AP condition holds for different users and disjointing cluster composition is not mandatory. This condition is related to the no-negligible contribution in terms of performance experienced by different users and that specif AP. In this context, the practical benefits are:

- reduced fronthaul signalling when only a subset of the APs must receive the downlink data intended for the UE and send their corresponding estimates of the uplink data to the CPU;
- reduced computational complexity when each AP needs to process signals of only a subset of the UEs.

3.2.1 System Model for Ul and Dl

To define a proper system model for the Cf-maMiMo, we need to consider the *coherence time-frequency block* $T_c \times B_c$, represented in fig.3.2, where T_c is a coherence time, that is the portion of time in which the channel is time-invariant, and B_c is the coherence bandwidth that is a frequency range in which the channel is constant also know as the frequency-flat condition. In this way, the channel between two antennas can be described by only one scalar coefficient and also if we assume the random realization in each coherence block, then the systems can be studied one block at a time without loss of generality (*block-fading model*). A signal fitting into this block

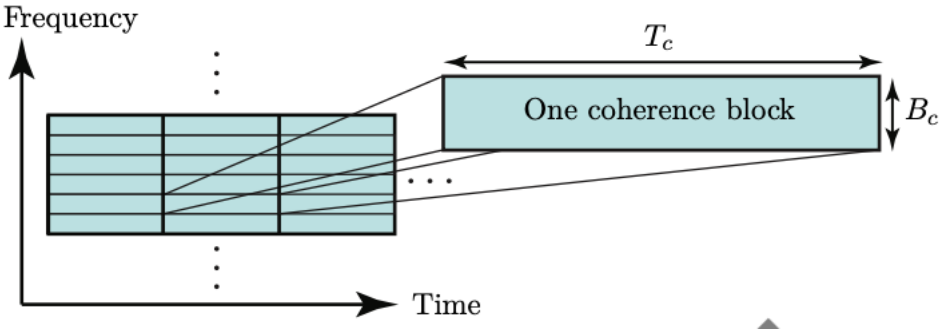


Figure 3.2: Cf-maMiMo $T_c \times B_c$ coherence block from [81]

can be described by $N = T_c \times B_c$ complex-valued samples. These parameters can be also called *transmission symbols* or simply *symbols*. A common approximation of the coherence time is $T_c = \lambda/(4\eta)$ where where λ is the wavelength and η is the velocity of the UE instead for $T_c = 1/(2T_d)$ where T_d is the time difference between the earliest and last propagation path (*channel delay spread*). Also we consider Cellular Massive MIMO TDD protocol. This one allows the APs to estimate the channel realizations, according to block-fading model and channel reciprocity. In particular, to estimate the channel realization between the APs and all users, the APs needs only the Ul pilots transmission [76], [75], [81]. The N samples, then, are distributed among the phases of the systems that are Ul training (τ_p samples), Dl-transmission(τ_d samples) and Ul-transmission(τ_u samples) as in fig.3.3. Hence $N = \tau_p + \tau_u + \tau_d$. Moreover,

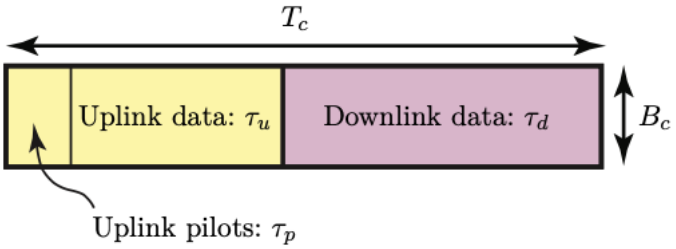


Figure 3.3: Cf-maMiMo $T_c \times B_c$ coherence block samples distribution from [81]

the most efficient way to estimate the channels is to consider a TDD protocol where each coherence block is used for both Ul and Dl transmissions. So we need only to transmit pilots only in the Ul and based on them each AP can then estimate the channels between itself and all the UEs. These channel estimates, then, can be utilized for both Ul and Dl. By applying coherent pre-coding in the Dl based on the channel estimates, the channel vector to each UE is transformed into a positive scalar channel. This scalar channel can be deduced at the UE from the Dl data signals,

without sending explicit pilots. According to the scenario depicted by the fig. 3.1 we define the channels between the AP m and users k as:

$$g_{mk} = \sqrt{\beta_{mk}} h_{mk} \tag{3.1}$$

where h_{mk} is the small-scale fading, and β_{mk} is the large-scale fading. Additionally h_{mk}, $m = \{1, ..., M\}$ and $k = \{1, ...K\}$, are independent and identically distributed (i.i.d.) RVs following $\mathcal{CN}(0, 1)$. As in [78], the justification of the assumption of independent small-scale fading is that the APs and the users are distributed over a wide area, and hence, the set of scatterers is likely to be different for each AP and each user. Given the coherence block's composition, during the training phase, the K users send simultaneously their τ_p-length pilot sequence to the APs., Let $\sqrt{\tau_p}\boldsymbol{\phi_k} \in \mathcal{C}^{\tau_p \times 1}$, $||\boldsymbol{\phi_k}||^2 = 1$ the sequence of the user k, we can define the received $\tau_p \times 1$ pilot vector from the AP m as follow:

$$\boldsymbol{Y}_m = \sqrt{\tau_p \rho_p} \sum_{k=1}^{K} g_{mk} \boldsymbol{\phi_k} + \boldsymbol{W}_m \tag{3.2}$$

where ρ_p is the normalized signal-to-noise ratio (SNR) of each pilot symbol and $\boldsymbol{W}_m$ is a vector of additive noise at the AP m. The elements of $\boldsymbol{W}_m$ are i.i.d.RVs following $\mathcal{CN}(0, 1)$. At the begin of the training phase, the AP m receive the pilot vector described in eq.ne 3.2. Then, it treats this vector as the real observation that it is going to use to derive the channel estimates $\hat{g_{mk}}$ for each users k it serves. Moreover, the channel estimates can be derived directly at AP or be delegated to CPU. In the last case the AP acts as relay and sends the pilot vector to the CPU. Since the channel vectors are independent, there isn't loss of the optimally if they are evaluated at the CPU. According to the definition of the coherence block and its distribution among phases, the τ_p samples allow the definition of τ_p orthogonal pilot sequences. This means that if in the system there are $K \geq \tau_p$ users, due to the pilot-sharing, those users will experience inter-user interference. So, given two pilot sequences k_1 and k_2 and according to TDD protocol, it even holds that

$$\boldsymbol{\phi}_{k_1}^H \boldsymbol{\phi}_{k_2} = \begin{cases} 1 \text{ if } k_1 = k_2 \in K_E \\ 0 \text{ otherwise} \end{cases} \tag{3.3}$$

So based on the pilot vector of eq.ne 3.2, to evaluate the channel estimates, the generic AP m needs to remove the interference by evaluating the projection of $\boldsymbol{Y}_m$ onto $\boldsymbol{\phi}_k^H$, that is $\boldsymbol{\phi}_k^H \boldsymbol{Y}_m$ for each user k. Let's

$$\tilde{y}_{mk} = \boldsymbol{\phi}_k^H \boldsymbol{Y}_m = \underbrace{\sqrt{\tau_p \rho_p} g_{mk}}_{Desired-Signal} + \underbrace{\sqrt{\tau_p \rho_p} \sum_{k'=1}^{K} g_{mk'} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}}_{Pilot-Contamination} + \underbrace{\boldsymbol{\phi_k} \boldsymbol{W}_m}_{Noise} \tag{3.4}$$

the evaluation of the projection. According to the 3.3, we can rearrange it as in the eq. ne 3.4 where the first term contains the desired channel scaled by a power coefficient we early described, while the second term represents the inter-user interference due to the pilot-sharing, also known as **Pilot-Contamination**, and then we have the noise term. At this point, we can apply the MMSE estimation process to retrieve the channel estimates as follow:

$$\hat{g}_{mk} = \frac{\mathrm{E}\{\tilde{y}_{mk} g_{mk}\}}{\mathrm{E}\{|\tilde{y}_{mk}|^2\}} \tilde{y}_{mk} = \theta_{mk} \tilde{y}_{mk} \tag{3.5}$$

with

$$\theta_{mk} \triangleq \frac{\sqrt{\tau_p \rho_p} \beta_{mk}}{\tau_p \rho_p \sum_{k'=1}^{K} \beta_{mk'} |\phi_k^H \phi_{k'}|^2 + 1} \tag{3.6}$$

It is worth noting that $\tau_p \geq K$, we can choose $\phi_1, ..., \phi_K$ pair-wisely orthogonal, and hence the *inter-user interference* term disappears. Then g_{mk} is independent from $g_{mk'}$ with $k \neq k'$. However, the limited length of the coherence interval, implies that holds the condition $\tau_p < K$, so mutually non-orthogonal pilot sequences must be used throughout the network. This implies that the channel estimates $\hat{g_{mk}}$ is degraded due to the pilot-sharing interference that is expressed by the second term of the eq. 3.4. The channel estimation error is also defined as $\tilde{g}_{mk} \triangleq g_{mk} - \hat{g}_{mk}$. Since we use MMSE estimator, $\hat{g}_{mk}$ is uncorrelated with $\tilde{g}_{mk}$. Since g_{mk} is Gaussian distributed, as described in 3.5, $\hat{g}_{mk}$ is the MMSE estimate of g_{mk}. This gives in turn that $\hat{g}_{mk}$ and $\tilde{g}_{mk}$ are independent.

In the **Dl-transmission phase**, the APs treat the channel estimates $\hat{g_{mk}}$ as the true channels, and use conjugate beam-forming to transmit signals to the K users. The signal can be defined as follow:

$$\boldsymbol{x}_m = \sqrt{\rho_{dl}} \sum_{k=1}^{K} \sqrt{\rho_{mk}} \hat{g}_{mk}^H s_k \tag{3.7}$$

with $\mathrm{E}\{|\boldsymbol{x}_m|^2\} \leq \rho_{dl}$ and $\mathrm{E}\{|s_k|^2\} = 1$. Moreover according to the channel model of eq. 3.1, the previous constraint can be rewritten as: $\sum_{k=1}^{K} \rho_{mk} \gamma_{mk} \leq 1$ for all m where

$$\gamma_{mk} = \mathrm{E}\{|\hat{g}_{mk}|^2\} = \sqrt{\tau_p \rho_p} \beta_{mk} \theta_{mk} \tag{3.8}$$

The generic user k will receive the following signal:

$$\begin{aligned}\boldsymbol{r}_k &= \sum_{m=1}^{M} g_{mk}\boldsymbol{x}_m + w_k = \\ &= \sum_{m=1}^{M} g_{mk}\sqrt{\rho_{dl}}\sum_{k'=1}^{K}\sqrt{\rho_{mk'}}\hat{g_{mk'}}^{H}s_{k'} + w_k = \\ &= \sqrt{\rho_{dl}}\sum_{m=1}^{M}\sum_{k'=1}^{K}\sqrt{\rho_{mk'}}\hat{g_{mk'}}^{H}g_{mk}s_{k'} + w_k\end{aligned} \tag{3.9}$$

where w_k is additive noise following $\mathcal{CN}(0,1)$. Then, the user will detect s_k from r_k. As in the work [76], we can re-arrange the signal of eq.ne 3.9 as follow:

$$\begin{aligned}\boldsymbol{r}_k &= \sum_{m=1}^{M} g_{mk}\boldsymbol{x}_m + w_k = \sqrt{\rho_{dl}}\sum_{m=1}^{M}\sum_{k'\neq k,k'=1}^{K}\sqrt{\rho_{mk'}}g_{m\hat{k'},k'}{}^{H}g_{mk}s_{k'} + w_k = \\ &= \underbrace{\sqrt{\rho_{dl}}\sum_{m=1}^{M}\sqrt{\rho_{mk}}\hat{g_{mk}}^{H}g_{mk}}_{DS_k}\,s_k + \underbrace{\sqrt{\rho_{dl}}\sum_{m=1}^{M}\sum_{k'\neq k,k'=1}^{K}\sqrt{\rho_{mk'}}\hat{g_{mk'}}^{H}g_{mk}}_{UI_{k,k'}}\,s_{k'} + \underbrace{w_k}_{N_k} = \\ &= DS_k + UI_{k,k'} + N_k\end{aligned} \tag{3.10}$$

From the eq.ne 3.10, we can derive the formula for the spectral efficiency $\boldsymbol{SE}_k$ as follow:

$$\boldsymbol{SE}_k^{Dl} = \log\left(1 + \frac{|DS_k|^2}{\mathrm{E}\{|UI_{k,k'}|^2\} + 1}\right) \tag{3.11}$$

To get the achievable data rate for Dl transmission for each user k we can use the expression: $\boldsymbol{R}_k^{Dl} = B_c \times \tau_d \times \boldsymbol{SE}_k^{Dl}$ according to [125].

In the **Ul-transmission phase** the K users send simultaneously their data to the APs. Before sending the data, each k user weights its symbol s_k, $\mathrm{E}\{|s_k|^2\} = 1$,with the power control coefficient $\sqrt{\rho_k}$ and $0 \leq \rho_k \leq 1$. So, the received signal during the Ul-transmission can be defined as follow:

$$\boldsymbol{y}_m = \sqrt{\rho_{ul}}\sum_{k=1}^{K} g_{mk}\sqrt{\rho_k}s_k + w_m \tag{3.12}$$

where ρ_{ul} is the normalized uplink SNR and w_m is additive noise at the AP m. Also here w_m are i.i.d.RVs following $\mathcal{CN}(0,1)$. To detect the transmitted symbol s_k the AP m multiplies the received signal $\boldsymbol{y}_m$ with the conjugate of its (locally obtained)

channel estimate $g_{\hat{mk}}$. Then, the obtained quantity $\hat{g}_{mk}^H \boldsymbol{y}_m$ is sent to the CPU via a backhaul network. The CPU will detect the symbol s_k from:

$$\begin{aligned}\boldsymbol{r}_k = \hat{\boldsymbol{g}}_{mk}^H \boldsymbol{y}_m = \sum_{m=1}^{M}\sum_{k'=1}^{K} \hat{g}_{mk}^H (g_{mk'}\sqrt{\rho_{ul}\rho_{k'}}s'_k + w_m) = \\ = \sum_{m=1}^{M}\sum_{k'=1}^{K} (\sqrt{\rho_{ul}\rho_{k'}}\hat{g}_{mk}^H g_{mk'}s'_k + \hat{g}_{mk}^H w_m) = \\ = \sum_{m=1}^{M}\sum_{k'=1}^{K} \sqrt{\rho_{ul}\rho_{k'}}\hat{g}_{mk}^H g_{mk'}s'_k + \sum_{m=1}^{M} \hat{g}_{mk}^H w_m\end{aligned} \tag{3.13}$$

where:

$$\hat{\boldsymbol{g}}_{mk}^H = \begin{bmatrix} \hat{g}_{1,k}^H \\ \hat{g}_{2,k}^H \\ \dots \\ \hat{g}_{M,k}^H \end{bmatrix} \tag{3.14}$$

is the combining vector for the users k. As in 3.10 and considering that the channel can be defined as the sum of the channel estimates $\hat{g}_{mk}$ and the channel estimation error $\tilde{g}_{mk}$,that is $g_{mk} = \tilde{g}_{mk} + \hat{g}_{mk}$, 3.13, also according to [81] can be re-arranged as follow:

$$\begin{aligned}\boldsymbol{r}_k = \sum_{m=1}^{M} \sqrt{\rho_{ul}\rho_k}\hat{g}_{mk}^H g_{mk}s_k + \sum_{m=1}^{M} \hat{g}_{mk}^H w_m = \\ = \sum_{m=1}^{M} \sqrt{\rho_{ul}\rho_k}\hat{g}_{mk}^H g_{mk}s_k + \sum_{m=1}^{M}\sum_{k\neq k',k'=1}^{K} \sqrt{\rho_{ul}\rho_{k'}}\hat{g}_{mk}^H g_{mk'}s'_k + \sum_{m=1}^{M} \hat{g}_{mk}^H w_m = \\ = \underbrace{\sum_{m=1}^{M} \sqrt{\rho_{ul}\rho_k}\hat{g}_{mk}^H \hat{g}_{mk}}_{DS_{\hat{g}}} s_k + \underbrace{\sum_{m=1}^{M} \sqrt{\rho_{ul}\rho_k}\hat{g}_{mk}^H \tilde{g}_{mk}}_{DS_{\tilde{g}}} s_k + \underbrace{\sum_{m=1}^{M}\sum_{k\neq k',k'=1}^{K} \sqrt{\rho_{ul}\rho_{k'}}\hat{g}_{mk}^H g_{mk'}}_{UI_{k,k'}} s'_k + \\ + \underbrace{\sum_{m=1}^{M} \hat{g}_{mk}^H w_m}_{N} =\end{aligned} \tag{3.15}$$

Where $DS_{\hat{g}}$ is the desired signal over estimated channel that can be used to detect the Ul transmitted data, $DS_{\tilde{g}}$ is the desired signal over unknown channel that can be considered as additional interference term, $UI_{k,k'}$ is the inter-users interference and N is the noise term. the $\boldsymbol{SE}_k^{Ul}$ can be derived as follow:

$$\boldsymbol{SE}_k^U = \mathbb{E}\{log_2(1 + \frac{|DS_{\hat{g}}|^2}{\mathbb{E}\{|DS_{\tilde{g}}|^2\} + \mathbb{E}|UI_{k,k'}|^2 + \mathbb{E}\{|N|^2\}})\} \tag{3.16}$$

Given $\boldsymbol{SE}_k^U$ for a specific user k, its data rate can be derived as follow $\boldsymbol{R}_k^{Ul} = B_c \times \tau_u \times \boldsymbol{SE}_k^{Ul}$.

3.2.2 Resources Allocation: Pilot Assignment Algorithms and Power Control/Allocation Schemes

In this section, we survey the most used techniques for the allocation of resources in a cell-free massive MiMo system. The main resources are Pilot sequences and Power. Hence, in the following section, we are going to survey Pilot Assignment Algorithms and Power Allocation Schemes. Generally, in a transmission system, the channel state information plays a vital role to provide a good level of performance. That information is usually acquired through pilot transmission in the **Ul pilot training phase**. Moreover, in the real system, the condition of having orthogonal pair-wise pilot sequences compels the UEs to share one pilot sequence at least. This phenomenon, which is also known as **Pilot Contamination**, reduces the channel estimation quality and makes the transmission less effective and harder to reject interference between those users sharing that pilot sequence. Thus, a properly designed pilot assignment procedure is critical to have good performance in the Cf-maMiMo system. Many pilot assignment procedures have been studied. The *Random Assignment Procedure* has been introduced in [126]. Here, each UE is randomly assigned a fixed pilot from the orthogonal pool and uses this pilot during the entire transmission. The drawback of this procedure is that neighbouring users occasionally could share the same pilot sequences generating mutual interference hard to suppress. The next algorithm, which has been introduced in [76] and [127], is based on the greedy approach. It assigns to the users the pilot sequence that minimizes the **Pilot contamination**. The authors in [76] and [127] asses that greedy-based pilot assignment procedures can converge to local optimum value but are unlikely to provide a globally optimal pilot assignment. With the work [128], user-center clustering has been introduced. Here the UEs are clustered into groups based on different large-scale information e.g.large-scale coefficients, UEs and APs location and distance. The clustering-based approach has been widely studied in [129], [130], [131]. Moreover in [130] and [131] the authors model the interference as graph. Here, the vertexes, representing UEs, are connected if at least an AP serves them. The objective of these algorithms is to colour the UEs with the fewest colours. The final assignment is achieved by updating the interference graph. The authors of [132] propose an iterative approach based on the Hungarian algorithm. With this algorithm, each UE and its neighbouring UEs are assigned with mutual orthogonal pilots following the *Hungarian algorithm*. When the performance

measures reach convergence or the iterations reach the allowed maximum number, the algorithm provides the final pilot assignment. All the aforementioned algorithms present limits in the sense that they might not be feasible for practical implementation because their complexity grows polynomially with the dimensions of the cell-free system, that is the number of UEs and APs. In other terms, they suffer from a lack of scalability. This consideration led the network engineers to the development of scalable algorithms as in [133], [134], [135]. Specifically, in [133], each UE selects its AP master, then each master AP informs a limited set of neighboring APs that it is going to serve that UE on specific assigned pilot. The neighbouring APs can decide whether join in serving that user or not according to its status. The scalability is achieved by providing the UE with the least bad pilot without performing optimization to provide a fairly pilot assignment. The authors in [134] and [135] used the K-means based approach for user clustering. The clusters are based on location and interference relationship between the cell-free actors UEs and APs. The UEs in the same cluster are assigned with orthogonal pilots. The approach in [134] divide the network into subareas, however it doesn't prevent the neighbouring UEs belonging to different cluster from pilot sharing. This issue has been solved by the authors in [135]. Additional information about pilot assignment procedures can be found in the survey [136]. According to the literature, the general consideration is that the algorithms with more complicated processing mechanisms and employing more information on cell-free actors and communication environment conditions will offer better performance. However, these kinds of algorithms are either unscalable or heuristic. This forces the network engineers to investigate further alternatives to introduce scalability by the exploitation of machine learning, thanks to its powerful signal processing ability in a heavily loaded network condition or user-centric clustering-based approach for the optimized pilot assignment or based ones. In a cell-free massive MiMo system, to have good performance, the available resources have to be efficiently handled. This holds even for power. To be more precise, given the K users in the system, they have to select appropriate transmit power for the Ul transmission, while the APs must allocate transmit power for Dl transmission. The procedure related to the power resource in a cell-free massive MiMo system is called power control for the UL-transmission and power allocation for DL-transmission. Generally, the power allocation/control procedure is regulated by a system-wide utility function that has to be optimized. The structure of the utility function implies the type of optimization problem that has to be solved and so the approach to solving it. According to the literature, the

most common types of optimization problems are max-min fairness, max sum SE, and max EE.

3.2.2.1 Max-Min fairness

The goal is to maximize the lowest SE among all the UEs in the network, which leads to uniform service, while the channel conditions will determine how good that service quality is. Since the SE of UE k is an increasing function of the effective SINR, maximizing the lowest SE is equivalent to maximizing the lowest effective SINR among all the UEs. There are several instances of the max-min fairness problem that can be shown to be convex or quasi-convex. Thus the optimal solution can be obtained by exploiting the bisection search, and the convex optimization as in [137], Geometric Programming as in [78] [79]. Moreover, there are instances of the max-min fairness problem that are non-convex, in which case one can sometimes find a local optimum by alternating optimization, which partitions the optimization variables into several sets and cyclically optimizes one at a time while keeping the other variable sets fixed [138] [139].

Although the aforementioned algorithms optimize the transmit power for all UEs to maximize the lowest SE in a system-wide manner, it is unavoidable that their computational complexities grow unboundedly with the number of the UEs, K, which makes these algorithms unscalable. Hence, distributed and heuristic schemes are needed to obtain scalable power control in large, practically implementable CF networks. Each device makes a local decision with limited involvement from the other devices. Fractional power control is a classical heuristic scheme in Ul multi-user systems. The principle of fractional power control is controlling the UE transmit power to compensate for a fraction of the path-loss differences among the UEs that are partially served by the same APs, where UE k selects its Ul to transmit power as in [81] and [134]

$$\rho_k = \frac{min_{i \in S_k} (\sum_{i \in M_i} \beta_{il})^v}{\sum_{k \in M_k} \beta_{kl}} \tag{3.17}$$

where $\sum_{k \in M_k} \beta_{kl}$ denotes the total channel gain from user k to the APs that serve it. $v \in [0, 1]$ indicate the power control behaviour. Specifically, $v = 0$ means that all the users transmit at the maximum power (full power transmission). A larger value of v implies each user to compensate for the variations in the total channel gain among the users in $i \in S_k$, which promotes more fairness. Fraction power allocation can

be used in the downlink, in which case AP l selects the downlink power allocation coefficient for user k proportionally to the channel gain, β_{kl}, as in [81]

$$\rho_{kl} = \frac{(\beta_{il})^v}{\sum_{i \in D_l} \beta_{kl}^v} \tag{3.18}$$

if $k \in D_l$ otherwise $\rho_{kl} = 0$. The exponent v implies the power allocation behaviour. A larger value of v gives a higher emphasis to the UEs according to their respective channel gains. D_l represent the set of the served users by the access point l.The $v = 0$ implies to allocating more power to the UEs in better channel conditions, which seems to be contrary to the max-min fairness concept. Apart from traditional optimization and heuristic methods, ML can be utilized to design power control/allocation methods. Such an approach cannot provide a better solution than the one found by classical optimization methods, but it could potentially lower the computational performance as in [140] There are ML-based schemes proposed to solve the max-min fairness problem in Cf-maMiMo as in [141]. Here, the authors proposed to approximately solve the max-min fairness problem using the local CSI by training a neural network to identify a mapping between that local CSI and the optimal solution to the system-wide max-min fairness problem.

3.2.2.2 Max Sum Spectral Efficiency (Max Sum SE)

A side effect of the max-min SE fairness problem is that a few UEs could degrade the overall system performance with bad channel conditions. However, in a large network is highly likely that the majority of the users reach larger SEs while affecting the users in the worst conditions. This is because the UE only causes interference to a small subset of neighbouring UEs. This consideration led the network engineers to consider the maximization of the sum SE representing the overall SE performance of the network instead of the SE achieved by a specific UE. The max sum SE problem is usually not convex. So it is impossible to obtain the optimal solution and therefore we need to settle for a local optimal value. The Successive Convex Approximation (SCA) is a promising methodology to handle the no-convexity of the "Max Sum SE" problem. As you can see in [142] [143] [144] SCA employs convex optimization where the non-convex term is substituted by its convex approximation. The aforementioned methods search for a local optimum due to the no-convexity of the problem. This allows the utilization of ML-based schemes because they have the potential of finding a better solution. Examples of this ML-based methods are [145] [146]. using Artificial Neural Networks (ANNs) In particular in [145] the authors use an artificial Neural

Network taking UE positions as input and the power control policy as output. Instead, the authors in [146] proposed a deep neural network-based power control method.

3.2.2.3 Max Energy Efficiency (Max EE)

The designing of large Cf-maMiMo network can be oriented on finding out "how much energy it takes to reliably transmit a certain amount of information". So the evaluation of energy efficiency is considered a good performance indicator of the system. Usually the energy efficiency is defined as:

$$EE = B\frac{\sum_{k=1}^{K} SE_k}{P_{total}} \tag{3.19}$$

where the B is the bandwidth and P_{total} is the total power consumption including four terms: the transmit power, a term accounting for the analogue processing of the transceiver chains, a term accounting for the digital signal processing, and a term for the fronthaul connections. Even this problem is a no-convex one. Hence, the SCA technique can provide local optimum value. In [147] the authors considered the EE maximization problem in an millimeter waves (mmWaves) Cf-maMiMo system. Here, they show that the problem is a no-convex. Moreover, they proposed a successive power-bound maximization method, that due to the merging of alternating optimization and sequential convex programming, provides the optimized values for the transmit power of each AP while keeping the transmit power of other APs fixed. As can be observed in [148], second-order methods performed very well, however they don't scale with the network size. This led network engineers to develop a first-order method to solve the "Max EE" problem. As you see in [149], the authors proposed a first-order method for the non-convex programming to the EE maximization problem that achieves the same performance with less run time than the second-order methods.

3.3 Multiplexing URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity

3.3.1 System Model

We consider the cell free massive MIMO downlink scenario represented in **Fig. 3.4**, with M APs, equipped with a single antenna, serving K single-antenna users. We assume a transmission scheme based on *conjugate beamforming*, which is one of the most prominent forms of linear precoding. Conjugate beamforming is based on a

pre-coding matrix proportional to the conjugate of the estimated channel matrix. In the time-domain, it is also known as "time-reversal beamforming," because it is equivalent to take the convolution of each transmitted symbols' sequence with its respective conjugated and time-reversed impulse-response estimate, and sum over the K correlations. Its reverse link counterpart, at the received side, is the matched-filtering.

The users are randomly distributed in a large area. This system supports two types of services, namely eMBB and URLLC, using network slicing and cell free massive MIMO. We model the network slicing by associating the contemporary multiple radio data streams of URLLC and eMBB users in the access network to one of the two partitions of the pool of $K = K_U \cup K_E$ users, with $K_U = \{1, 2, \ldots, i, \ldots, h\}$ and $K_E = \{h+1, \ldots, h+i, \ldots, K\}$, which are linked to the set of APs $M = M_U \cup M_E$ partitioned in the two subsets $M_U = \{1, 2, \ldots, m_u, \ldots, |M_U|\}$ and $M_E = \{1, \ldots, m_e, \ldots, |M_E|\}$. The RAN slices associated to the two considered services are defined as:

$$eMBB_{slice} \stackrel{\text{def}}{=} (M_E, K_E) \tag{3.20}$$

$$URLLC_{slice} \stackrel{\text{def}}{=} (M_U, K_U) \tag{3.21}$$

3.3.2 URLLC and eMBB Multiplexing by Network Slicing Cf-maMiMo

We assume a cell-free massive MIMO-OFDM system as in [150], where in downlink (DL) transmissions the $N_{sc} = 2400$ subcarriers are divided into multiple resource blocks (RB), which are allocated to UEs such that the specific service requirements

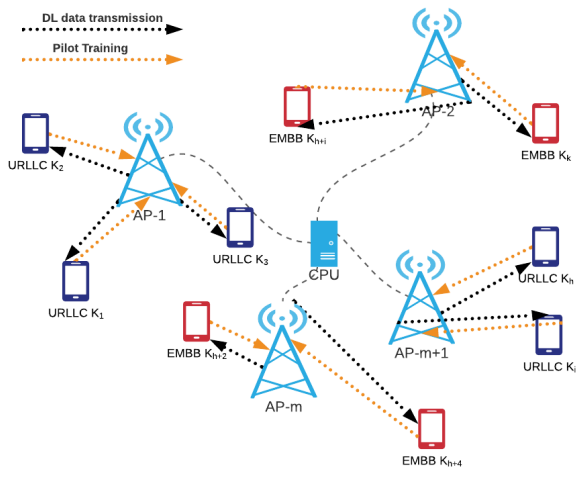


Figure 3.4: Cf-maMiMo downlink scenario.

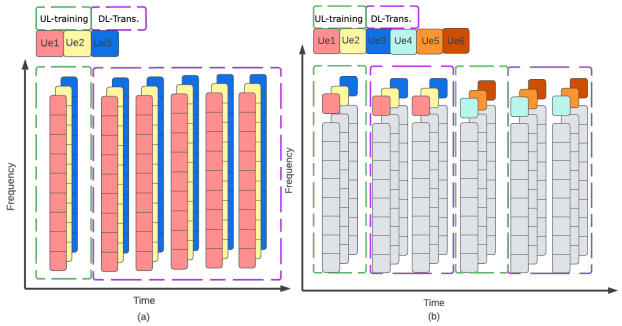


Figure 3.5: Resource block allocation scheme in a coherence interval: (a) each UE occupies all RBs for eMBB; (b) each UE occupies only one RB for URLLC.

are fulfilled. According to the 5G new radio (NR) standard, we assign 12 subcarriers to a RB, such that the bandwidth of each RB is $12\Delta f$, with $\Delta f = 7.5$ kHz.

In the time domain, according to the massive MIMO operation, the subcarriers and time samples are grouped to fit the *coherence interval* (CI), which consists of a number of subcarriers and time samples over which the channel response can be approximated as constant and flat-fading. If the coherence bandwidth is B_c and the coherence time is T_c, then each CI contains $N = B_c \cdot T_c$ complex-valued samples[1].

Fig. 3.5 shows the RB allocation scheme in a CI, where we distinguish the CSI acquisition phase, labeled by "UL-training," and the data transmission phase, labeled by "DL-Trans." Actually, the total number of N samples is distributed among the training phase, which uses N_p samples, and the DL and UL transmissions, which use N_d and N_u samples, respectively, such that $N = N_p + N_d + N_u$ [81]. In this work we focus only on the training phase and downlink transmission, hence we have $N_u = 0$.

To multiplex the two considered heterogeneous services, namely the eMBB and URLLC, we adopt the resource allocation scheme shown in Fig. 3.5. Since the eMBB is designed to meet the demand for high-data rate transmission, each UE occupies all subcarriers, i.e., each UE is enabled to transmit with maximum bandwidth. For the URLLC case, we assume that multiple UEs can be scheduled for transmission in the same coherence interval, as the size of URLLC packets is much smaller than that of eMBB. Theoretically, assuming that each RB serves k URLLC UEs, the cell-free massive MIMO-OFDM system can support a total of $kN_{sc}/12$ URLLC UEs transmitting simultaneously [150].

[1] A common rule of thumb is the following:

- $T_c = 1/(4 \cdot v)$ with v equal to user's velocity;
- $B_c = 1/(2T_d)$ where T_d is the delay difference between longest and shortest path.

We formalize the RB allocation scheme described above by assuming that the URLLC utilizes a fraction δ_B of the coherence bandwidth B_c and a fraction of the downlink transmission time, as explained in Sec. 3.3.6. This model reflects in the calculation of the data rates for the two services.

In the following, we analyze the training phase and propose the best suitable scheme for coexisting URLLC and eMBB services (Sec. 3.3.3); we also present two pilot assignment schemes, one simply random and the other based on a greedy selection (in Sec. 3.3.4).

3.3.3 Training phase

The application of network slicing to cell free massive MIMO imposes a re-planning of the training phase. By design, the training phase is performed in **uplink**. Let's $N_p^{(E)}$ and $N_p^{(U)}$ be the number of samples assigned to the training phase of eMBB and URLLC services, respectively.

According to the strategies provided in [151], we distinguish the following three cases, shown in **Fig. 3.6**, which differ in the applied configurations for the CIs training and transmission phases:

1. **Independent Training**: URLLC and EMBB users are **independently** trained in **different** CIs followed by its own DL transmission in the same CI. This means that the full set of N samples available in a CI are assigned either to eMBB (i.e., $N = N^{(E)}$) or to URLLC (i.e., $N = N^{(U)}$) users during the full duration of a CI. Being $N = N_p + N_d$, the samples for the training phase are $N_p^{(E)}$ and $N_p^{(U)}$ and the samples for the DL transmissions are given by either $N_d^{(E)} = N - N_p^{(E)}$ or $N_d^{(U)} = N - N_p^{(U)}$, respectively.

2. **Contemporary Training**: URLLC and eMBB users are **contemporarily** trained in the **same** CI, followed by contemporary DL transmissions in the same CI. This means that $N_p = \max\{N_p^{(E)}, N_p^{(U)}\}$ for both services. In other terms, for both services, the system uses N_p samples for the UL training phase and the remaining $N_d = N - N_p \leq N - \max\{N_p^{(E)}, N_p^{(U)}\}$ for DL transmission. Also, $N_p = \max\{N_p^{(E)}, N_p^{(U)}\}$ guarantees orthogonality between users but it is not sufficient to prevent the interference between services.

3. **Sequential Training**: URLLC and eMBB users are **sequentially** trained in the **same** CI, followed by contemporary DL transmissions in the same CI. Hence, we have $N_p = N_p^{(E)} + N_p^{(U)} \leq 2 \cdot \max\{N_p^{(E)}, N_p^{(U)}\}$, with the constraint

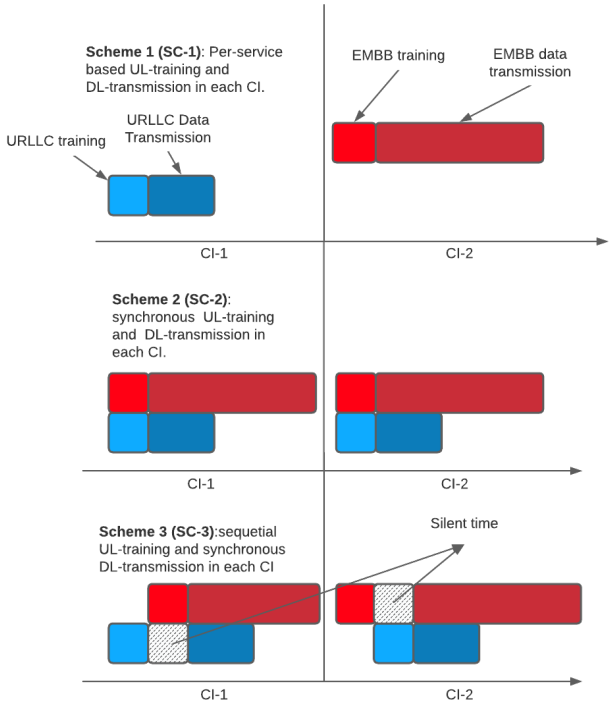


Figure 3.6: Possible UL-training and DL-transmission schemes

that one type of users is silent while the other performs the training phase. Therefore, the DL accounts a number of samples $N_d = N - N_p \leq N - 2 \cdot \max\{N_p^{(E)}, N_p^{(U)}\}$.

To define *independent* pilot sequences using $N_p^{(U)}$ training samples for URLLC users and $N_p^{(E)}$ training samples for eMBB users, we assume that

$$N_p^{(U)} = |K_U| \tag{3.22}$$

$$N_p^{(E)} = |K_E|. \tag{3.23}$$

The conditions (3.22) and (3.23) imply that user-coherent interference doesn't exist. However, the orthogonality is not ensured between P_U and P_E, giving rise to *inter-service interference*.

Depending on the applied training phase configuration, under the assumption of per service pairwise orthogonality of the sequence pilot sets, we have:

1. For **independent training**, URLLC and eMBB users will not experience any kind of inter-users or inter-service interference. Hence, based on the type of users which is active in the current CI, the signal received by the m-th AP during the training phase will by either: $\mathbf{y}_m = \mathbf{U} + \mathbf{w}_m$ or $\mathbf{y}_m = \mathbf{E} + \mathbf{w}_m$ for all N_p

samples in the CI for each of the two services, where the three contributions are related respectively to the training of the K_U URLLC users, i.e., $\mathbf{U} \in \mathbb{C}^{N_p^{(U)} \times 1}$, and the K_E eMBB users, i.e., $\mathbf{E} \in \mathbb{C}^{N_p^{(E)} \times 1}$, and, finally, to the noise term $\mathbf{w}_m \in \mathbb{C}^{\max\{N_p^{(E)}, N_p^{(U)}\} \times 1}$. The elements of the three terms are all i.i.d random variables distributed according to a complex normal distribution, i.e., $\sim \mathcal{CN}(0, 1)$. It is important to observe that the sum of the complex vectors $\mathbf{U}$, $\mathbf{E}$, $\mathbf{w}_m$ is defined for equal-size vector spaces.

2. For **contemporary training**, URLLC and eMBB users will experience only *inter-service interference* and the signal received by the m-th AP is

$$\mathbf{y}_m = \mathbf{U} + \mathbf{E} + \mathbf{w}_m \tag{3.24}$$

for $N_p = \max\{N_p^{(E)}, N_p^{(U)}\}$ samples in the CI.

3. For **sequential training**, URLLC and eMBB users will not experience any kind of user or inter-service interference, as in the first case, and similarly the signal received by m-th AP results either $\mathbf{y}_m = \mathbf{U} + \mathbf{w}_m$ or $\mathbf{y}_m = \mathbf{E} + \mathbf{w}_m$, respectively for the $N_p^{(U)}$ and $N_p^{(E)}$ samples in the CI.

Despite the independent training and the sequential training provide no interference among both users and services, it presents significant drawbacks. In the first case of independent training, as shown in Fig. 3.6, each available CI is assigned to the sole transmission of either URLLC or eMBB, with subsequent under-utilization of the available resources (more evident in the case of URLLC transmissions, based on a very short payload), and increase of URLLC latency. Hence this scheme is not suitable to fulfil URLLC latency requirements.

Also the third scheme, i.e. the sequential training, is not suitable because it violates both reliability and latency requirements for the URLLC, being the training done in the absence of interference; instead, the transmissions are simultaneous to eMBB traffic, but rely in a pilot assignment based on non-accurate channel estimations, which impacts the reliability of URLLC transmissions. Moreover, the sequential training implies the introduction of a delay that increases the URLLC latency.

Hence, for the optimal coexistence of URLLC and eMBB services, due to their specific QoS requirements, the only possible scheme is the contemporary trainings of URLLC and eMBB users in the same CI. This is the scheme adopted throughout the subsequent analyses presented in this paper.

Based on the type of service, i.e., URLLC and eMBB, we define two sets of pilot sequence signals P_E and P_U as:

$$P_U = \{\sqrt{N_p^{(U)}\rho_u}\boldsymbol{\varphi}_i^{(U)} \in \mathbb{C}^{N_p^{(U)}\times 1} | i \in K_U\} \tag{3.25}$$

$$P_E = \{\sqrt{N_p^{(E)}\rho_e}\boldsymbol{\varphi}_j^{(E)} \in \mathbb{C}^{N_p^{(E)}\times 1} | j \in K_E\} \tag{3.26}$$

where

$$\{\sqrt{N_p^{(U)}\rho_u}\boldsymbol{\Phi}^{(U)}\} \in \mathbb{C}^{N_p^{(U)}\times |K_U|}$$

and

$$\{\sqrt{N_p^{(E)}\rho_e}\boldsymbol{\Phi}^{(E)}\} \in \mathbb{C}^{N_p^{(E)}\times |K_E|}$$

are the sets $\boldsymbol{\Phi}^{(U)} = [\boldsymbol{\varphi}_1^{(U)} \dots \boldsymbol{\varphi}_i^{(U)} \dots \boldsymbol{\varphi}_{K_U}^{(U)}]$ and $\boldsymbol{\Phi}^{(E)} = [\boldsymbol{\varphi}_1^{(E)} \dots \boldsymbol{\varphi}_j^{(E)} \dots \boldsymbol{\varphi}_{K_E}^{(E)}]$ of $N_p^{(U)}$-length and $N_p^{(E)}$-length pilot signals to be assigned to URLLC and eMBB users, with ρ_u and ρ_e the UL pilot powers.

Hence, for the contemporary scheme, the general expression for the signal received by the m-th AP during the training phase of the massive MIMO conjugate beamforming scheme, when all K users transmit the pilots simultaneously, is given by (3.24) with

$$\begin{aligned} \mathbf{U} &= \sqrt{N_p^{(U)}\rho_{\mathrm{u}}} \sum_{i\in K_U} g_{mi}^{(U)} \boldsymbol{\varphi}_i^{(U)} \\ \mathbf{E} &= \sqrt{N_p^{(E)}\rho_{\mathrm{e}}} \sum_{i\in K_E} g_{mi}^{(E)} \boldsymbol{\varphi}_i^{(E)} \end{aligned} \tag{3.27}$$

The channel gains $g_{mi}^{(U)}$ and $g_{mi}^{(E)}$ are defined as [76]

$$g_{mi}^{(U)} = \sqrt{\beta_{mi}^{(U)}} h_{mi}^{(U)} \quad \text{for all } i \in K_U \tag{3.28}$$

$$g_{mi}^{(E)} = \sqrt{\beta_{mi}^{(E)}} h_{mi}^{(E)} \quad \text{for all } i \in K_E, \tag{3.29}$$

with the large scale fading coefficients $\beta_{mi}^{(E)}$ and $\beta_{mi}^{(U)}$, and the small scale fading coefficients $h_{mi}^{(U)} \sim \mathcal{CN}(0,1)$ and $h_{mi}^{(E)} \sim \mathcal{CN}(0,1)$, modeled as i.i.d. complex Gaussian RVs according to Sec. 2.5.2 of [81]. The assumption of independent small-scale fading is fully justified in the proposed scenario, because the APs and the users are distributed over a wide area, such that the set of scatterers is likely to be different for each pair AP-user.

Given the two sets of pilot sequence signals P_U and P_E defined in (3.25) and (3.26), for each of the two sets of users K_U and K_E, we can derive the following projections. For the user $i \in K_E$ the projection is given by:

$$\tilde{y}_{mi}^{(E)} \triangleq \boldsymbol{\varphi}_i^{(E)H} \mathbf{y}_m = \boldsymbol{\varphi}_i^{(E)H} \left[\mathbf{U} + \mathbf{E} + \mathbf{w}_m\right] \tag{3.30}$$

where $\mathbf{U}$ and $\mathbf{E}$ are given by (3.27).

Taking into account the assumption of the pair-wise orthogonality, i.e.:

$$\boldsymbol{\varphi}_i^{(U)H} \boldsymbol{\varphi}_t^{(U)} = \begin{cases} ||(\boldsymbol{\varphi}_t^{(U)})||^2 = \boldsymbol{I}_t \text{ if } t = i \in K_U \\ 0 \text{ otherwise} \end{cases}$$

the projection for the user $i \in K_E$ simplifies to:

$$\begin{aligned} \tilde{y}_{mi}^{(E)} &= \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} g_{mi}^{(E)} + \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} \sum_{s \in K_U} g_{ms}^{(U)} \boldsymbol{\varphi}_i^{(E)H} \boldsymbol{\varphi}_s^{(U)} \\ &+ \boldsymbol{\varphi}_i^{(E)H} \mathbf{w}_m \end{aligned} \tag{3.31}$$

Similarly, for the user $j \in K_U$ the projection is given by:

$$\begin{aligned} \tilde{y}_{mj}^{(U)} \triangleq \boldsymbol{\varphi}_j^{(U)H} \mathbf{y}_m &= \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{mj}^{(U)} + \\ &+ \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} \sum_{t \in K_E} g_{mt}^{(E)} \boldsymbol{\varphi}_j^{(U)H} \boldsymbol{\varphi}_t^{(E)} + \boldsymbol{\varphi}_j^{(U)H} \mathbf{w}_m \end{aligned} \tag{3.32}$$

given the pair-wise orthogonality condition in the eMBB case:

$$\boldsymbol{\varphi}_i^{(E)} H \boldsymbol{\varphi}_s^{(E)} = \begin{cases} ||(\boldsymbol{\varphi}_s^{(E)})||^2 = \boldsymbol{I}_s \text{ if } s = i \in K_E \\ 0 \text{ otherwise} \end{cases}$$

The second term in both (3.31) and (3.32) represents the **network slicing pilot contamination**.

Based on the received projection $\tilde{y}_{mj}^{(U)}$ or $\tilde{y}_{mi}^{(E)}$, the CPU evaluates the MMSE estimates $\hat{g}_{mj}^{(U)}$ and $\hat{g}_{mi}^{(E)}$ for the users belonging to the sets K_U and K_E, needed to apply the precoding for the conjugate beamforming.

By assumption, the large scale fading coefficients $\hat{g}_{mj}^{(U)}$ are known to be complex-gaussian distributed, i.e., $\sim \mathcal{CN}(0, \beta_j^{(U)})$, hence the MMSE estimator for the URLLC case is

$$\hat{g}_{mj}^{(U)} = \frac{\mathbb{E}\left\{g_{mj}^{(U)} \tilde{y}_{mj}^{(U)}\right\}}{\mathbb{E}\left\{|\tilde{y}_{mj}^{(U)}|^2\right\}} \tilde{y}_{mj}^{(U)} = c_{mj}^{(U)} \tilde{y}_{mj}^{(U)} \tag{3.33}$$

where

$$c_{mj}^{(U)} = \frac{\sqrt{N_p^{(U)}\rho_u}\beta_{mj}^{(U)}}{1+\sqrt{N_p^{(E)}\rho_e}\sum_{t\in K_E}\beta_{mt}^{(E)}|\boldsymbol{\varphi}_j^{(U)H}\boldsymbol{\varphi}_t^{(E)}|^2} \tag{3.34}$$

and, as in [78], the mean square value $\gamma_{mi}^{(U)}$ is defined as:

$$\begin{aligned}\gamma_{mj}^{(U)} &= \mathbb{E}\left\{|\hat{g}_{mj}^{(U)}|^2\right\} = \sqrt{N_p^{(U)}\rho_u}\beta_{mj}^{(U)}c_{mj}^{(U)} = \\ &= \frac{N_p^{(U)}\rho_u(\beta_{mj}^{(U)})^2}{1+\sqrt{N_p^{(E)}\rho_e}\sum_{t\in K_E}\beta_{mt}^{(E)}|\boldsymbol{\varphi}_j^{(U)H}\boldsymbol{\varphi}_t^{(E)}|^2}\end{aligned} \tag{3.35}$$

Being the channels of the single-antennas APs statistically identical, $\gamma_{mj}^{(U)}$ is the same for all M_U APs [78].

The channel estimation error is defined as $\epsilon_j^{(U)} = \hat{g}_j^{(U)} - g_j^{(U)}$ where $\epsilon_j^{(U)} \sim \mathcal{CN}(0, \beta_j^{(U)} - \gamma_j^{(U)})$ is independent from $\hat{g}_j^{(U)} \sim \mathcal{CN}(0, \gamma_j^{(U)})$. The mean square error has a variance of $\mathbb{E}\{|\epsilon_j^{(U)}|^2\} = \beta_i^{(U)} - \gamma_j^{(U)}$ [81].

Similar derivations can be done to obtain the channel estimate $\hat{g}_{mi}^{(E)}$ for $i \in K_E$.

3.3.4 Pilot assignment procedures

We consider two PAs, customized for the specific scenario of Cf-maMiMo URLLC and eMBB sliced services.

The *Random PA* is a very simple assignment method constrained only by the user's type. With this strategy the user will randomly receive a pilot out from the pilots' set. Users of different services that are in close vicinity could experience inter-service interference which is not taken into account for the selection of the assigned pilots.

The *Greedy PA*, instead, accounts for the inter-service interference, i.e. for the *network slice pilot contamination*. The pilot contamination is expressed for the two services by the terms $\sum_{t\in K_E} g_{mt}^{(E)}\boldsymbol{\varphi}_i^{(U)H}\boldsymbol{\varphi}_t^{(E)}$ and $\sum_{s\in K_U} g_{ms}^{(U)}\boldsymbol{\varphi}_i^{(E)H}\boldsymbol{\varphi}_s^{(U)}$ in (3.31) and (3.32). Their variances are given by:

$$\mathbb{E}\left\{|\sum_{t\in K_E} g_{mt}^{(E)}\boldsymbol{\varphi}_i^{(U)H}\boldsymbol{\varphi}_t^{(E)}|^2\right\} = \sum_{t\in K_E}\beta_{mt}^{(E)}|\boldsymbol{\varphi}_i^{(U)H}\boldsymbol{\varphi}_t^{(E)}|^2$$

$$\mathbb{E}\left\{\left|\sum_{s\in K_U} g_{ms}^{(U)}\boldsymbol{\varphi}_i^{(E)H}\boldsymbol{\varphi}_s^{(U)}\right|^2\right\} = \sum_{s\in K_U}\beta_{ms}^{(U)}|\boldsymbol{\varphi}_i^{(E)H}\boldsymbol{\varphi}_s^{(U)}|^2$$

Each user will be assigned the pilot sequence which minimizes the network slice pilot contamination summed over all the APs, according to the type of service. For the user $j \in K_U$ the assigned pilot is found by solving (P1), for the user $i \in K_E$ by solving (P2):

$$\arg\min_{\boldsymbol{\varphi}_j^{(U)}} \sum_{m \in M_U} \sum_{\substack{t \in K_E, \\ \boldsymbol{\varphi}_j^{(U)} = \boldsymbol{\varphi}_i^{(E)}}} \beta_{mt}^{(E)} |\boldsymbol{\varphi}_j^{(U)H} \boldsymbol{\varphi}_t^{(E)}|^2 \tag{P1}$$

$$\arg\min_{\boldsymbol{\varphi}_i^{(E)}} \sum_{m \in M_E} \sum_{\substack{s \in K_U, \\ \boldsymbol{\varphi}_i^{(E)} = \boldsymbol{\varphi}_s^{(U)}}} \beta_{ms}^{(U)} |\boldsymbol{\varphi}_i^{(E)H} \boldsymbol{\varphi}_s^{(U)}|^2 \tag{P2}$$

3.3.5 Downlink transmission

Let s_u, with $u \in K_U$ and $\mathbb{E}\{|s_u|^2\} = 1$, be the symbol to be transmitted to URLLC users, and s_e, with $e \in K_E$ and $\mathbb{E}\{|s_e|^2\} = 1$, the symbol to be transmitted to eMBB users. We assume that the symbols are uncorrelated, i.e., $\mathbb{E}\{|s_e s_u|^2\} = 0$. Based on the assumption of channel reciprocity for the single-antenna APs, the signals transmitted by the AP $m_u \in M_U$ and the AP $m_e \in M_E$, by applying the conjugate beamforming precoding based on the MMSE channel estimation, can be written as:

$$x_{m_u \in M_U} = \sum_{u \in K_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} \hat{g}_{m_u, u} s_u \tag{3.36}$$

$$x_{m_e \in M_E} = \sum_{e \in K_E} \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} \hat{g}_{m_e, e} s_e \tag{3.37}$$

with the conditions $\mathbb{E}\{|x_{m_u}|^2\} \leq \rho_{\mathrm{u}}$, $\mathbb{E}\{|x_{m_e}|^2\} \leq \rho_{\mathrm{e}}$.

In this Cf-maMiMo scenario, the generic i-th user receives a signal resulting from the superposition of (3.36) and (3.37). In particular, the signals received by the URLLC user $u \in K_U$ and the eMBB user $e \in K_E$ are expressed as (in the following we omit the superscripts $^{(U)}$ and $^{(E)}$ to improve the clarity of the expressions - the terms are distinguished by the subscripts):

$$
\begin{aligned}
r_u &= \sum_{m_u \in M_U} g_{m_u,u} x_{m_u} + \sum_{m_e \in M_E} g_{m_e,u} x_{m_e} + n_u = \\
&= \sum_{m_u \in M_U} \sum_{u' \in K_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u} \hat{g}_{m_u,u'} s_{u'} + \\
&\quad + \sum_{m_e \in M_E} \sum_{e \in K_E} \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} g_{m_e,u} \hat{g}_{m_e,e} s_e + n_u
\end{aligned} \tag{3.38}
$$

and

$$
\begin{aligned}
r_e &= \sum_{m_e \in M_E} g_{m_e,e} x_{m_e} + \sum_{m_u \in M_U} g_{m_u,e} x_{m_u} + n_e = \\
&\quad \sum_{m_e \in M_E} \sum_{e' \in K_E} \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} g_{m_e,e} \hat{g}_{m_e,e'} s_{e'} + \\
&\quad + \sum_{m_u \in M_U} \sum_{u \in K_U} \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} g_{m_u,e} \hat{g}_{m_u,u} s_u + n_e
\end{aligned} \tag{3.39}
$$

that can be expressed in a compact form [152] as:

$$
r_u = S_u s_u + T_u s_u + I_{u,u'} s_{u'} + I_{u,e} s_e + n_u \tag{3.40}
$$

where:

$$
\begin{aligned}
S_u &= \mathbb{E}\left\{ \sum_{m_u \in M_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u} \hat{g}_{m_u,u} \right\} \\
T_u &= \sum_{m_u \in M_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u} \hat{g}_{m_u,u} \\
&\quad - \mathbb{E}\left\{ \sum_{m_u \in M_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u} \hat{g}_{m_u,u} \right\} \\
I_{u,u'} &= \sum_{\substack{u' \in K_U \\ u' \neq u}} \sum_{m_u \in M_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u'} \hat{g}_{m_u,u'} \\
I_{u,e} &= \sum_{e \in K_E} \sum_{m_e \in M_E} \sqrt{N_p^{(E)} \rho_{\mathrm{e}}} g_{m_e,u} \hat{g}_{m_e,e}
\end{aligned}
$$

are, respectively, a deterministic factor S_u scaling the desired signal of the URLLC user $u \in K_U$ given by the conjugate beamforming precoding; the contribution T_u in the received signal related to the *beamforming uncertainty gain*; the interference $I_{u,u'}$ with the signals by other URLLC users $u' \in K_U \setminus \{u\}$, which is $I_{u,u'} = 0$ under

our assumption of orthogonality between users of the same service; the *inter-slice interference* $I_{u,e}$ between the URLLC user u and all eMBB users; and the effective noise n_u perceived by the user u.

3.3.6 Performance metrics

We define three metrics to measure the performance of the proposed Cf-maMiMo system with network slicing applied to URLLC and eMBB served users: the spectral efficiency, the achievable average data rate and the outage probability.

The DL spectral efficiency for users $u \in K_U$, under the condition $M >> K$, is defined as [152]:

$$\begin{aligned} SE_u &= \log_2 \left(1 + \frac{|S_u|^2}{\mathbb{E}\left\{|I_{u,u'}|^2\right\} + \mathbb{E}\left\{|I_{u,e}|^2\right\} + 1} \right) = \\ &= \log_2 \left(1 + \frac{|S_u|^2}{\mathbb{E}\left\{|I_{u,e}|^2\right\} + 1} \right) \end{aligned} \tag{3.41}$$

where, as derived in the Appendix 3.3.9:

$$|S_u|^2 = N_p^{(U)} \rho_{\mathrm{u}} \left(\sum_{m_u \in M_U} \gamma_{m_u,u} \right)^2 \tag{3.42}$$

and

$$\mathbb{E}\left\{|I_{u,e}|^2\right\} = N_p^{(E)} \rho_{\mathrm{e}} \sum_{e \in K_E} \sum_{m_e \in M_E} \gamma_{m_e,e} \beta_{m_e,u}. \tag{3.43}$$

The expression (3.41) holds under the condition $M >> K$, as demonstrated in [153], which analyses the achievable data rates based on the number of the APs for downlink transmission, showing that for large values of M the contribution of the beamforming uncertainty gain in the received signal is negligible.

For the user $e \in K_E$ we define the corresponding terms S_e, $I_{e,e'}$ and $I_{e,u}$ and derive the expression for the spectral efficiency. However, in light of the adopted RB allocation strategy, since the eMBB payload has a bigger size than the URLLC, we assume that URLLC DL transmissions are "puncturing" the DL time slot, resulting in the utilization of a portion of the total DL time T_d.

The DL time slot is given by $T_d = T_c(1 - N_p/N)$, with $N = B_c T_c$. We identify two portions in T_d: the *contention-based DL transmission time* defined as $T_{cb} = w_{cb} T_d$, where simultaneous URLLC and eMBB transmissions cause inter-service interference; and the *contention-free DL transmission time* $T_{cf} = T_d - T_{cb}$, where only eMBB users transmit and will not experience any kind of inter-user or inter-slice interference. The

weighting factor $w_{cb} \in (0, \frac{1}{2})$ represents the estimated time fraction of simultaneous DL transmissions.

Hence, the total spectral efficiency for eMBB users is:

$$SE_e = SE_e^{cb} + SE_e^{cf} \tag{3.44}$$

where

$$\begin{aligned} SE_e^{cb} &= \log_2\left(1 + \frac{|S_e|^2}{\mathbb{E}\{|I_{e,e'}|^2\} + \mathbb{E}\{|I_{e,u}|^2\} + 1}\right) \\ &= \log_2\left(1 + \frac{|S_e|^2}{\mathbb{E}\{|I_{e,u}|^2\} + 1}\right) \\ SE_e^{cf} &= \log_2\left(1 + \frac{|S_e|^2}{\mathbb{E}\{|I_{e,e'}|^2\} + \mathbb{E}\{|I_{e,u}|^2\} + 1}\right) \\ &= \log_2\left(1 + |S_e|^2\right) \end{aligned}$$

are, respectively, the *contention-based eMBB spectral efficiency term* SE_e^{cb} related to T_{cb} and the *contention-free eMBB spectral efficiency term* related to T_{cf}.

Therefore, we may define the achievable data rate for DL transmission of user $u \in K_u$, R_u^{DL}, and user $e \in K_e$, R_e^{DL}, as functions of the spectral efficiency SE_u and SE_e, the coherence bandwidth B_c, the *contention-based DL transmission time* T_{cb} and the *contention-free DL transmission time* T_{cf} as:

$$R_u^{DL} = \delta_B \cdot B_c \cdot T_{cb} \cdot SE_u \tag{3.45}$$

$$R_e^{DL} = B_c \cdot \{T_{cb} \cdot SE_e^{cb} + T_{cf} \cdot SE_e^{cf}\} \tag{3.46}$$

where δ_B is the fraction of the coherence bandwidth assigned to URLLC users according to our RB assignment scheme.

Finally, we define the outage probability as a metric to evaluate the *scalability* of Cf-maMiMo supporting network slicing. The eMBB and URLLC services have specific payload size requirements b_u and b_e, respectively. The outage probability refers to a network operating condition that doesn't support the transmission of packets of these fixed sizes of b_u and b_e bits. Then, if $R_u^{th} = f(b_u)$ and $R_e^{th} = f(b_e)$ are the data rate threshold values for the URLLC and eMBB services,

the outage probability for the specific user is defined as

$$P_u^{out} = \Pr\left[R_u^{DL} < R_u^{th}\right] \text{ for } u \in K_U \tag{3.47}$$

$$P_e^{out} = \Pr\left[R_e^{DL} < R_e^{th}\right] \text{ for } e \in K_E \tag{3.48}$$

3.3.7 Performance evaluation

We investigate the performance of the proposed network sliced Cf-maMiMo system for different sizes of the pool of users, that is $20 \leq K \leq 200$, still fulfilling the design condition of cell-free massive MiMo, $M >> K$, being $M = 200$ the total number of APs in the area with $M = M_E + M_U$ and $|M_E| = |M_U|$.

3.3.7.1 Simulation Setup

We consider a square area of 1 km^2 with M APs and K users, with $M_U = M/2$ APs and $K_U = K/2$ users (UEs) assigned to URLLC and $M_E = M/2$ APs and $K_E = K/2$ UEs assigned to eMBB. To avoid boundary effects and to imitate a network spread over an infinite area, the square is wrapped around. Regardless of service, during the UL training phase we assume a maximum transmission power of $\rho_u^{tot} = \rho_e^{tot} = 100\ mW$ for each user. For the DL transmission, the available maximum total power is $\rho_u^{tot} = \rho_e^{tot} = 200\ mW$. The APs apply a centralized power allocation scheme and assign to each users an amount of power given by

$$\rho_{mk}^u = \rho_u^{tot} \frac{(\beta_{km}^u)^v}{\sum_{(k,m)} (\beta_{km}^u)^v}$$

with $k \in K_U$, $m \in M_U$ and $v \in [0, 1]$. The value of $v = 0$ implies that the assigned power is a fraction of the total available power. The value of $v = 1$ implies that the assigned power relies on the channel condition between the access point m and the considered user k. Similarly, for the eMBB APs and users. According to [81], we assume that the length of a CI is $200 \leq N \leq O(10^4)$, where $N = 200$ corresponds to a network with high mobility and high channel dispersion, whereas $N = O(10^4)$ reflects a network with low user mobility and low channel dispersion. We set $N = 5 * 10^4$. The APs are deployed in a urban dense micro-cell scenario matching the 3GPP model for the 2 GHz band. According to that model, for each m-AP and k-UE, we evaluate the channel gains as:

$$\beta_{km}[dB] = -30.5 - 36.7 \log_{10}(d_{km}/1m) + F_{km} \tag{3.49}$$

where the d_{km} is the 3D distance from the m-th AP of k-th UE and $F_{km} \sim \mathcal{N}(0, 4^2)$ is the shadowing term. The shadowing fading from the pair m-th AP and k-th UE to the pair j-th AP and i-th UE has a correlation given by:

$$\mathbb{E}[F_{km} F_{ij}] = \begin{cases} 4^2 2^{\Delta_{ki}/9\ m} \text{ if } m = j \\ 0 \text{ if } m \neq j \end{cases} \tag{3.50}$$

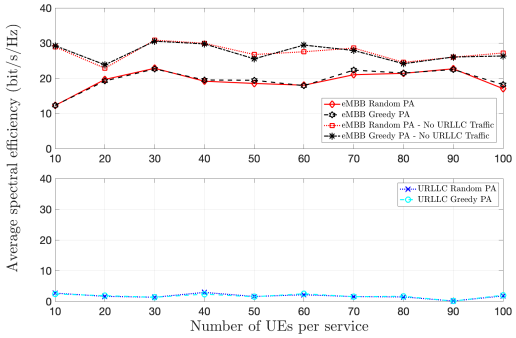


Figure 3.7: Spectral efficiency of the two services URLLC (lower plot) and eMBB (upper plot) multiplexed in the same area in the CfmaMiMo system for the two considered PAs. For the eMBB, the spectral efficiency is shown for both scenarios with and without URLLC traffic.

where Δ_{ki} is the distance of user k from user i. The second row reflects the shadowing effects for different APs that is assumed to be zero. The reason is that APs are typically distributed in the network at a distance larger than tens of meters and purposely deployed to view the deployment area from different directions than other APs. The coherence bandwidth is assumed to be $Bc = 1$ MHz. Finally, the eMBB packet payload size is set to $b_e = 1500$ bytes and the URLLC size to $b_u = 32$ bytes, according to the standard.

3.3.7.2 Results and Discussions

We show the average spectral efficiencies and the cumulative distribution functions (CDFs) of the throughput of URLLC and eMBB users for the two different pilot assignment schemes against the number of users. To show the benefits of the spatial diversity offered by the Cf-maMiMo, we analyze also the case of eMBB traffic only, without any interference from URLLC users. The average spectral efficiency vs. the number of UEs per service is depicted in Fig. 3.7 for the eMBB with and without URLLC traffic (upper plot) and for the URLLC (lower plot). As expected, in a single-service deployment scenario, the spectral efficiency is higher with respect to the multi-service scenario and this reflects in a higher throughput of the eMBB users in the absence of URLLC traffic, as shown in Fig. 3.8 for the two exemplary cases of 40 and 100 UEs. Apparently, the PA scheme does not have any impact on the performance, but this point deserves further investigations.

The Fig.s 3.8 and 3.9 represent the CDFs of the throughput for the eMBB and the URLLC services, respectively, for the two considered pilot assignment schemes

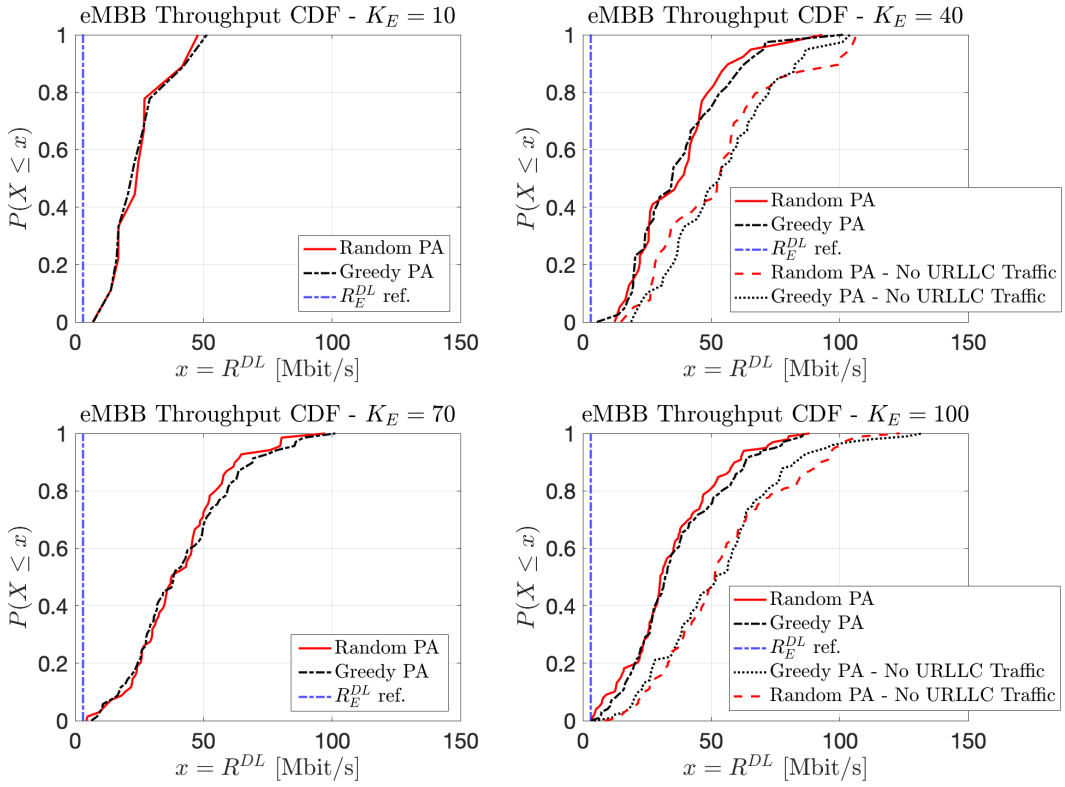


Figure 3.8: CDFs of the eMBB throughput for different numbers of UEs for the two considered PAs.

in the proposed cell-free massive MiMo radio access network. We notice that the throughput does not show any improvement when the more sophisticated greedy PA scheme is applied. Increasing the number of users from 10 to 100 per service, while keeping constant the number of APs, $M = 200$, produces at first an improvement of the throughput (from 10 to 40 UEs), but then a slight performance downgrade in the case of multiplexed eMBB and URLLC. This is related to the main condition of the cell free massive MiMo asset, that is $M >> K$, i.e. the number of APs must be much greater than the number of served users. Indeed this effect is not evident in the eMBB CDFs without URLLC traffic, because in this case the number of APs is twice the number of UEs. When M approaches the number of UEs in the area, the spectral efficiency expression (3.41) should account also for the beamforming uncertainty gain.

For the URLLC, we observe a similar behavior, with the throughput re-scaled according to the URLLC traffic characteristics. Moreover, Fig. 3.9 shows that for low levels of interference, i.e. for small numbers of users (i.e., 10 or 40), the greedy PA scheme slightly outperforms the random PA; instead, at high interference regimes,

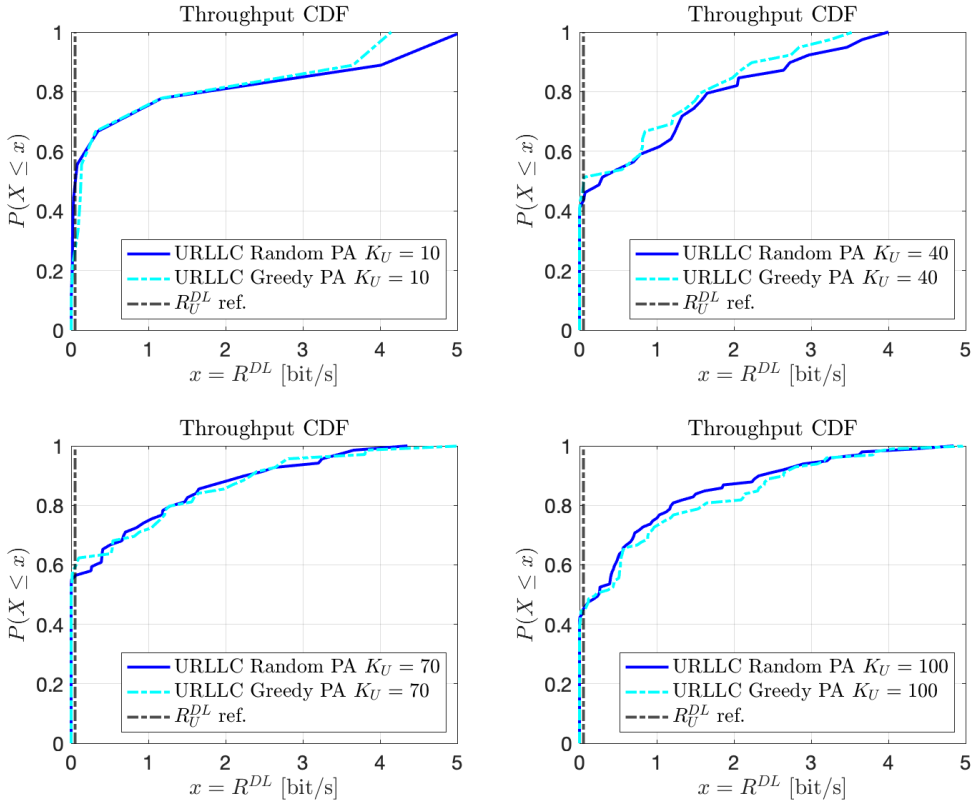


Figure 3.9: CDFs of the URLLC throughput for different numbers of UEs for the two considered PAs.

i.e. with 70 or 100 users per service, the adopted PA scheme does not have any effect.

Finally, looking at the intersection between the throughput CDFs curves and the reference values per service R_u^{th} and R_e^{th} reported in Fig.s 3.8 and 3.9, we may infer that the service slicing in the proposed Cf-maMiMo RAN guarantees the fulfillment of the QoS requirements for eMBB users with a probability of 100%, whereas for the URLLC the outage probability remains at acceptable levels.

3.3.8 Conclusion and future directions

In this work we studied the coexistence of URLLC and eMBB service on a cell free massive MIMO RAN. In particular, we devised a new theoretical model accounting for new assumptions related to network slicing applied to the CfmaMiMo to multiplex the heterogeneous services.

We introduced two customized procedures to assign pilot sequences to users, namely the Random PA and Greedy PA, where the latter has been defined with

the purpose of minimizing the network slice pilot contamination. The results show, however, that in the proposed configuration the PA choice does not influence the performance except in a low interference regime. This point deserves further investigations.

Future directions of this work include the performance evaluation of the same system with different propagation conditions and different precoding techniques, also relaxing the assumption of the orthogonality between the two sets of pilot sequences. Finally, we plan to study alternative pilot procedure assignments, including implementations based on reinforcement learning algorithms.

3.3.9 Derivation of the SE terms

In the following we report the detailed calculations of the terms S_u and $I_{u,e}$ of the closed-form expression (3.41) for the spectral efficiency. We remind that for a user i served by the m-th AP the channel estimation error is $\epsilon_{mi} = \hat{g_{mi}} - g_{mi}$, and, as in [81], the following conditions hold: $\mathrm{E}\{\epsilon_{mi}\hat{g_{mi}}\} = 0$ and $\mathrm{E}\{\epsilon_{mi}(\tilde{y}_{mi}^{(U)})^H\} = 0$. This is due to the fact that the channel estimation error ϵ_{mi} and the estimates $\hat{g_{mi}}$ are jointly Gaussian distributed, thus they are uncorrelated. For the S_u term in the SE expression we have:

$$\begin{aligned} |S_u|^2 &= \left| \mathbb{E}\left\{ \sum_{m_u \in M_U} \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} g_{m_u,u} \hat{g}_{m_u,u} \right\} \right|^2 \\ &= \left| \sqrt{N_p^{(U)} \rho_{\mathrm{u}}} \mathbb{E}\left\{ \sum_{m_u \in M_U} g_{m_u,u} \hat{g}_{m_u,u} \right\} \right|^2 . \end{aligned} \tag{3.51}$$

By expressing $g_{m_u,u}$ in terms of the estimation error $\epsilon_{m_u,u}$, we have:

$$\begin{aligned} &\mathbb{E}\left\{ \sum_{m_u \in M_U} g_{m_u,u} \hat{g}_{m_u,u} \right\} = \\ &= \mathbb{E}\left\{ \sum_{m_u \in M_U} \left(\hat{g}_{m_u,u} - \epsilon_{m_u,u} \right) \hat{g}_{m_u,u} \right\} = \\ &= \mathbb{E}\left\{ \sum_{m_u \in M_U} \left(\hat{g}_{m_u,u} \hat{g}_{m_u,u} - \epsilon_{m_u,u} \hat{g}_{m_u,u} \right) \right\} = \\ &= \mathbb{E}\left\{ \sum_{m_u \in M_U} |\hat{g}_{m_u,u}|^2 - \sum_{m_u \in M_U} \epsilon_{m_u,u} \hat{g}_{m_u,u} \right\} = \\ &= \sum_{m_u \in M_U} \mathbb{E}\left\{ |\hat{g}_{m_u,u}|^2 \right\} - \sum_{m_u \in M_U} \mathbb{E}\left\{ \epsilon_{m_u,u} \hat{g}_{m_u,u} \right\} \end{aligned} \tag{3.52}$$

With the condition $\mathrm{E}\{\epsilon_{mi}\hat{g_{mi}}\} = 0$, we have

$$\begin{aligned} |S_u|^2 &= \left| \sqrt{N_p^{(U)}\rho_{\mathrm{u}}} \sum_{m_u \in M_U} \gamma_{m_u,u} \right|^2 = \\ &= N_p^{(U)}\rho_{\mathrm{u}} \left(\sum_{m_u \in M_U} \gamma_{m_u,u} \right)^2 \end{aligned} \tag{3.53}$$

where $\gamma_{(m_u,u)}$ is defined in (3.35).

With the same considerations, we can obtain the term $I_{u,e}$:

$$\begin{aligned} \mathbb{E}\left\{|I_{u,e}|^2\right\} &= \left| \mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} \sqrt{N_p^{(E)}\rho_{\mathrm{e}}} g_{m_e,u}\hat{g}_{m_e,e} \right\} \right|^2 = \\ &= \left| \sqrt{N_p^{(E)}\rho_{\mathrm{e}}} \mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} g_{m_e,u}\hat{g}_{m_e,e} \right\} \right|^2 . \end{aligned} \tag{3.54}$$

By expressing $g_{m_e,u}$ in terms of the estimation error $\epsilon_{m_e,u}$:

$$\begin{aligned} &\mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} g_{m_e,u}\hat{g}_{m_e,e} \right\} = \\ &= \mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} \left(\hat{g}_{m_e,u} - \epsilon_{m_e,u}\right) \hat{g}_{m_e,e} \right\} . \end{aligned} \tag{3.55}$$

Defining

$$G_{\mathrm{U}E} = \mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} \hat{g}_{m_e,u}\hat{g}_{m_e,e} \right\} \tag{3.56}$$

and

$$C_{\epsilon\hat{g}} = \mathbb{E}\left\{ \sum_{e \in K_E} \sum_{m_e \in M_E} \epsilon_{m_e,u}\hat{g}_{m_e,e} \right\}, \tag{3.57}$$

we have

$$\begin{aligned} |I_{u,e}|^2 &= N_p^{(E)}\rho_{\mathrm{e}} \left|G_{\mathrm{U}E} - C_{\epsilon\hat{g}}\right|^2 = \\ &= N_p^{(E)}\rho_{\mathrm{e}} \left(G_{\mathrm{U}E}^2 + C_{\epsilon\hat{g}}^2 - 2G_{\mathrm{U}E}C_{\epsilon\hat{g}}\right) . \end{aligned} \tag{3.58}$$

With some algebra, we have:

$$\begin{aligned} G_{\mathrm{U}E}^2 &= \sum_{e \in K_E} \sum_{m_e \in M_E} \mathbb{E}\left\{ \left|\hat{g}_{(m_e,u)}\right|^2 \right\} \mathbb{E}\left\{ |\hat{g}_{m_e,e}|^2 \right\} = \\ &= \sum_{e \in K_E} \sum_{m_e \in M_E} \gamma_{m_e,u}\gamma_{m_e,e}, \end{aligned} \tag{3.59}$$

$$
\begin{aligned}
C^2_{\epsilon\hat{g}} &= \sum_{e\in K_E}\sum_{m_e\in M_E} \mathbb{E}\left\{|\epsilon_{m_e,u}|^2\right\}\mathbb{E}\left\{|\hat{g}_{m_e,e}|^2\right\} = \\
&= \sum_{e\in K_E}\sum_{m_e\in M_E} \left(\beta_{m_e,u}-\gamma_{m_e,u}\right)\gamma_{m_e,e}.
\end{aligned}
\tag{3.60}
$$

With the condition $\mathrm{E}\{\epsilon_{mi}\hat{g_{mi}}\}=0$, we have

$$
\begin{aligned}
&G_{\mathrm{UE}}C_{\epsilon\hat{g}} = \\
&= \sum_{e\in K_E}\sum_{m_e\in M_E} \mathbb{E}\left\{\epsilon_{m_e,u}\hat{g}_{m_e,u}\right\}\mathbb{E}\left\{\hat{g}_{m_e,e}\hat{g}_{m_e,e}\right\} = 0
\end{aligned}
\tag{3.61}
$$

and (3.58) simplifies to

$$
\begin{aligned}
|I_{u,e}|^2 &= N_p^{(E)}\rho_{\mathrm{e}}\sum_{e\in K_E}\sum_{m_e\in M_E}\left[\gamma_{m_e,e}\cdot\right. \\
&\quad \left.\cdot\left(\gamma_{m_e,u}+\beta_{m_e,u}-\gamma_{m_e,u}\right)\right] = \\
&= N_p^{(E)}\rho_{\mathrm{e}}\sum_{e\in K_E}\sum_{m_e\in M_E}\gamma_{m_e,e}\beta_{m_e,u}.
\end{aligned}
\tag{3.62}
$$

3.4 Summary

In this chapter , we introduced the cell-free massive mimo radio access network topology together with the motivation which pushed researches and engineers to the infield deployment. The benefits, such as huge and constant data rate or high coverage, coming from its capabilities, allow the cell free massive mimo of outperforming the legacy systems. This also justifies its huge deployment in different contexts. The main aspects of this topology such as training, uplink and downlink phases have been presented and used in the presented research work titled **Multiplexing URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity**. In this work, we devised a new theoretical model accounting our network slicing view with the purpose of multiplexing multiple services, then we compared customized procedures handling the assignment of the resources and discovered that in low interference regime the performance are not affected by the chosen procedure.

Chapter 4

Artificial Intelligence-Supported Resource management

In this section, we introduce the Reinforcement Learning approach. This kind of approach belongs to the field of machine learning. Due to the capability of finding optimal solutions for network problems even with partial knowledge of the network scenario that the model is observing, the RL-based models are the most promising modelling methodology to overcome complex modelling issues of network problems. To understand the reason behind this huge potential we arrange this chapter as follows: in sec: **4.1 Reinforcement Learning**, we introduce the reinforcement learning methodology including its mathematical fundamentals. In particular, in sec **4.1.1 Markov Decision Process, Markov Game, Mean Field Game**, we present the Markov Decision Process and its alternative models following the configuration of the observed environment. These alternatives are Markov Game and the Mean field game. Moreover, in sec: **4.1.2 Value and Q-value functions, optimal policy, Q-learning**, we introduce the concept of value functions, Q-value function, optimal policy and the basic Q-learning technique. In sec: **4.1.3 Value-based techniques** and **4.1.4 Policy-based techniques**, we introduce the most used technique to solve the RL-based problems. Then in the sec: **4.2 From Reinforcement to Deep Reinforcement Learning** we explain why the classical Reinforcement Learning (RL)-based modelling design process of the network problem may shift to the Deep Reinforcement Learning (DRL)-based one. Then in the sec: **4.3 DRL-based Radio Resource Management**, we provide a list of interesting works touching on the covered topic of the DRL applied to radio resource management. Instead, in the sec: **4.4 Example of RL based-modelling applied to the presented problems**, we provide basic ideas related to the application of RL-framework on the presented problems. Those ideas focus on the definitions of the fundamental elements of the

MDP process that eventually can be used to develop fully RL-based projects related to the presented topics.

4.1 Reinforcement Learning

In this section, we present the basic framework of Reinforcement learning. This framework is known in the RL-based literature as the "Agent-Environment" interacting framework or simpler RL-environment. As in fig.4.1, we have an environment

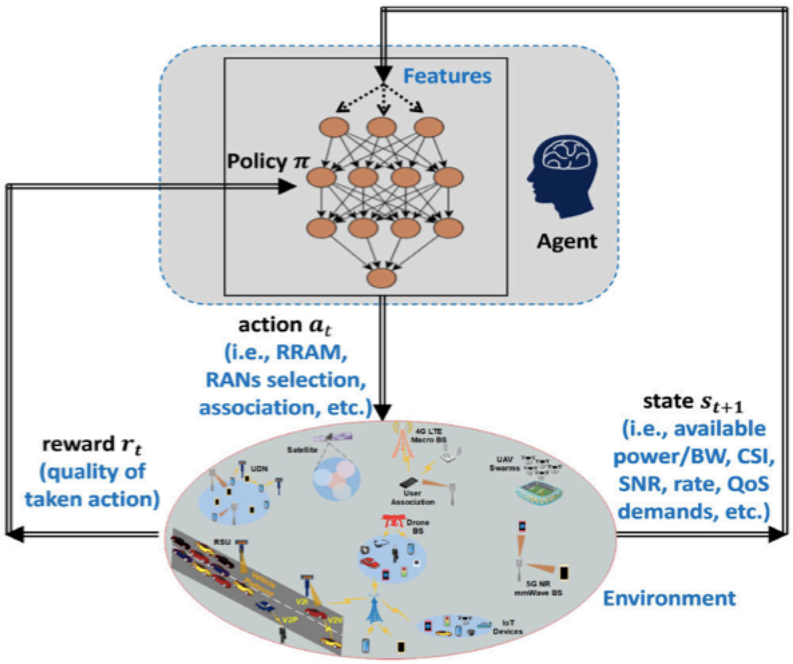


Figure 4.1: Reinforcement Learning Environment [154]

that is observed by an agent. According to the current environment's state, the Agent selects an action that will be applied to the environment. This will produce a state shift towards the next stage in the environment. From the Agent's point of view, this is a decision-making control loop that takes in input a state and according to the inner criteria of the decision-making function of the agents, produces, as output, an action that will be applied to the observed environment. The control-loop execution is regulated by specific functions that can evaluate the goodness of the executed action and its implied state. Precisely, according to the fig.4.2, the task of the RL is to calculate a **behaviour strategy**, also known as **policy**, with the purpose of **maximizing a satisfaction criterion**. In detail, this Agent-Environment framework consists of a decision-maker, called the "Agent", operating in an environment represented by the state s_t. The agent can take certain action a_t , as a consequence of

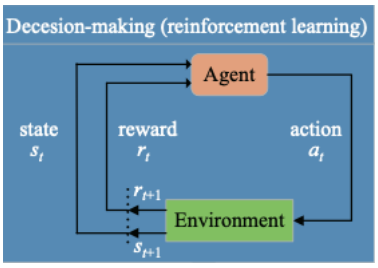


Figure 4.2: Explanation of Reinforcement learning working-flow [154]

the evaluation of the current state s_t . After choosing an action a_t time t, the agent receives reward r_{t+1} and finds itself in a new state s_{t+1} depending on the current state and the chosen action. The mathematical foundations of RL are based on presented Markov Decision Process (MDP) in 4.1.1. Moreover, in sec.4.1.1, we introduce also the alternatives of the MDP that are known as the Markov game and Mean Field Game. Also, we introduce the concept of value functions and optimal policy.

4.1.1 Markov Decision Process, Markov Game, Mean Field Game

The **MDP** is the basic describing formalism of RL environment. In particular, it is a discrete-time stochastic control process. The MDP can be described by a tuple (S, A, p, r), where S is a the states set, A is the actions set, $p : (S \times A) \to (S)$ is a transition probability function giving the probability to reach state s_{t+1} from state s_t and the execution of action a_t. The $r : (S \times A) \to (\mathbb{R})$ is the reward function giving the reward r_t that is the obtained immediate reward after action a_t is performed. The fig.4.3 describe the MDP's fundamental property. It means that the future of

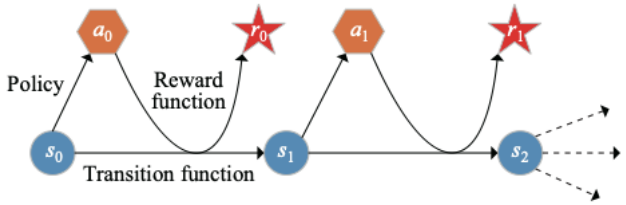


Figure 4.3: Markov decision process description [154]

the process only depends on the current state, and the agent has no interest in the full history. Formally the probability of being in the state s_{t+1} depends only by the previous state s_t and the performed action a_t. It can be described as follow:

$$p(s_{t+1}|s_0, a_0, ..., s_t, a_t) = p(s_{t+1}|s_t, a_t) \tag{4.1}$$

where p is the transition probability function earlier described. As we said before, the task of the agent in the RL-the based framework is to find the best behaviour strategy or policy maximizing a satisfaction criterion. Formally, let's this policy π a function mapping a state s with action a, that is $\pi : (S) \rightarrow (A)$. The goal of the MDP consists in finding the optimal policy π^* that maximizes the reward function. The MDP can be finite or infinite time horizon. Due to these features, the objective function changes. In particular,

$$\pi^* = argmax_\pi \begin{cases} \mathbb{E}\{\sum_{t=0}^{T} \gamma r(s_t, \pi(s_t))\} \text{ finite time horizon T} \\ \lim_{T\rightarrow\infty} inf \mathbb{E}\{\sum_{t=0}^{T} \gamma r(s_t, \pi(s_t))\} \text{ infinite time horizon T} \end{cases} \tag{4.2}$$

where $\gamma \in [0, 1]$ allows the definition of discounted values for the rewards. The γ factor defines how important the future rewards respect the current one. In other terms, if $\gamma = 0$, the agent is said to be "myopic" because it tries to maximise always the current rewards. In contrast, if $\gamma \rightarrow 1$, the agent will strive for a long-term higher reward. With MDPs, generally, we assume that the system's state is fully observable

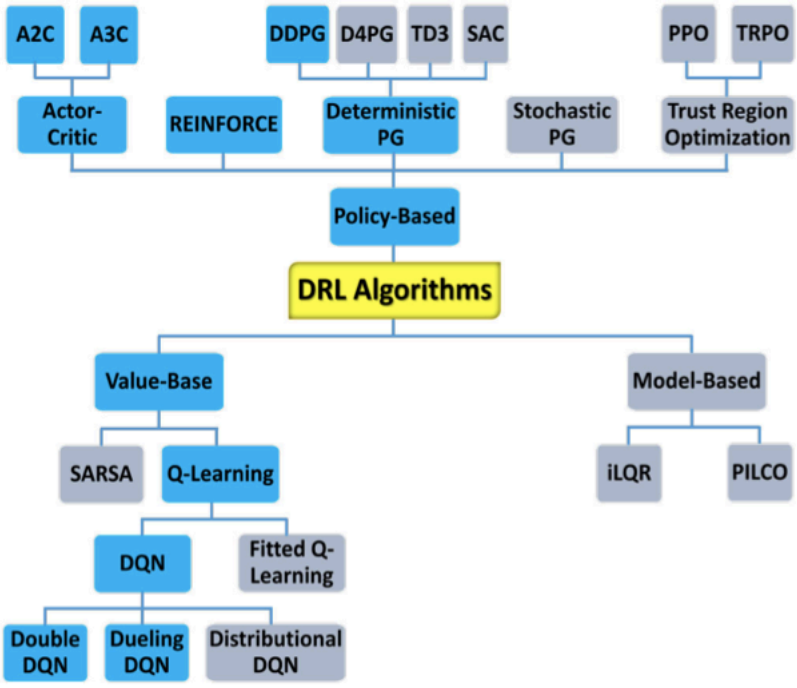


Figure 4.4: Deep Reinforcement Learning Taxonomy [154]

by the agent. However, in many cases, the agent only observes a partial state of it. This is the case of a complex scenario where the formalisation of the whole state of the environment is very complex or impossible. However, the MDP with

partially observed state, that is called Partially Observed Markov Decision Process (POMDP), can be altered with the introduction of observation of the system and observation probability function. In this case, the POMDP can be formalised as the tuple $(S, A, p, r, \Omega, \mathcal{O})$. The difference respect with to the classical MDP definition is the presence of $\Omega, \mathcal{O}$ components. Precisely $\mathcal{O} : (\Omega \times S \times A \times S) \rightarrow [0, 1]$ is a function returning the probability that the agent obtains the observation $o \in \Omega$[1] after action $a_t \in A$ is taken at state $s_t \in S$ and the agent moves to state $s_{t+1} \in S$.At each iteration, the agent being in the state s_t, selects an action a_t based on its belief about the current state s_t, i.e. $b(s_t)$, and observes the immediate reward r_t and current observation o. Based on the observation o and its belief about the current state $b(s_t)$, the agent, then, updates its belief about the new state s_{t+1}, i.e., $b(s_{t+1})$. The update is regulated according to [155] by the following:

$$b(s_{t+1}) = \frac{\mathcal{O}(o|s_t, a_t, s_{t+1}) \sum_{s \in S} p(s_{t+1}|s_t, a_t) b(s_t)}{\sum_{s_{t+1} \in \Omega} \mathcal{O}(o|s_t, a_t, s_{t+1}) \sum_{s \in S} p(s_{t+1}|s_t, a_t) b(s_t)} \tag{4.3}$$

. Similarly to MDP model, the agent in POMDP aims to find the optimal policy π to maximize its expected long-term discounted reward defined as in 4.2. Since the radio access network is a complex network segment, it is reasonable to think that a single agent monitoring the whole segment is not tractable. Therefore is preferable the definition of a multi-agent framework for solving network problems. To model a multi-agents framework, the basic MDP has to be converted either in a **Markov Game (MG)** or and **Mean Field Game (MFG)**. The **MG**, in game theory, is a dynamic game with probabilistic transitions played by multiple players/agents. The MG is defined as the tuple $(\mathcal{I}, S, \{A^i\}_{i \in \mathcal{I}}, p, \{r^i\}_{i \in \mathcal{I}})$ where

- $\mathcal{I} \triangleq \{1, ..., i, ...I\}$ set of players;
- $S \triangleq \{S^1, ..., S^i, ...S^I\}$ is the global states space set where S^i is the S^i's states space set.
- $\{A^i\}_{i \in \mathcal{I}}$ are the set of actions space and A^i is the action space set of agent i
- $\{r^i\}_{i \in \mathcal{I}}$ is the set of payoff functions defined as $r^i : (S \times A^1 \times ... \times A^I) \rightarrow \mathbb{R}$. The payoffs are obtained after the executions of all the actions of the agents.
- $p : (S \times A^1 \times ... \times A^I) \rightarrow [0, 1]$ is the transition probability function of the Markov game

[1] instead Ω is the set of observation

In the Markov Game, all the agents start at initial state s_0 and after the observation of the current state s_t, they will select an actions $a_t = \{a_t, ..., a_t^i, ...a_t^I\}$ and after the execution, they will receive the payoff r_t^i from the application of the reward function and the new observation. Contemporary, the system will transit to the new state s_{t+1} according to the probability $p(s_{t+1}|s_t, a_t)$. In this game the agents try to find their optimal policy $\mathit{\Pi} = \{\pi_1^*, ..., \pi_i^*, ...\pi_I^*\}$ that is called **Game's Equilibrium**. The equilibrium of the game corresponds to a situation where the expected long-run rewards functions provide the optimal values, that is $\mathbb{E}\{\sum_{t=0}^{\infty} \gamma r_t^i(s_t, \pi^*(s_t))\}$. When the number of the player and state set is finite the game equilibrium is called **Nash Equilibrium**. With an infinite state set, the reward payoff is the discounted one.[2]. In general, a N-stochastic games become hard to analyse when $N \to \infty$. In this case, the **MFG** is a useful framework to manage such games where there are $\mathcal{N} = \{1, 2, ..., N\}$ players. The basic idea of MFG is the following:

- Assume that $\mathcal{N}$ players are identical, indistinguishable and interchangeable (Homogeneous Hypothesis).
- Assume that the reward and the dynamic for each player are known.
- So, when $N \to \infty$ the global state S of the all players can be split in S_t^i and population state S_t^{-i} with population distribution $\mu_t(s_t^i) = \lim_{N \to \infty} \frac{\sum_{j \neq i, j=1}^{N} \mathbb{I}_{s_t^j = s_t^i}}{N}$. $\mathbb{I}_{s_t^j = s_t^i} = 1$ when $s_t^j = s_t^i$ and 0 otherwise.
- Due to the Homogeneous Hypothesis of the players, it is possible focus on representative player i and solve the problem of maximizing the expected long-run payoff $\mathbb{E}\{\sum_{t=0}^{\infty} \gamma r_t^i(s_t, \pi_t(s_t, \mu_t(s_t)), \mu_t(s_t))|s_0 = s\}$ with the constraint $s_{t+1} \sim \mathcal{P}(s_t, a_t, \mu_t)$

Hence, according to the main setting-idea of MFG, at time t, after the representative player chooses her action a_t according to some policy π_t, she will receive reward $r_t(s_t, a_t, \mu_t)$ and its state will evolve under a controlled stochastic dynamics of a mean-field type $P(.|s_t, a_t, \mu_t)$. Here, the policy π_t depends on both the current state s_t and the current population state distribution μ_t such that $\pi : \mathcal{P}(S) \to \mathcal{P}(A)$. However, in the classical MFG, the reward r and the dynamic for each player $\mathcal{P}$ are unknown. This allows us to cover the major part of possible scenarios. In that case, at time t, after the representative player i chooses the action a_t according to some policy $\pi : (S \times \mathcal{P}(S) \to \mathcal{P}(A)$, it will receive a reward $r_t(s_t, a_t, \mathcal{L}_t)$ and its state will

[2]Further details about the stochastic game can be found in [156]

evolve according to $P(.|s_t, a_t, \mathcal{L}_t)$. The objective of the player is to solve the following control problem $\mathbb{E}\{\sum_{t=0}^{\infty} \gamma r_t^i(s_t, \pi_t(s_t, \mu_t(s_t)), \mathcal{L}_t)|s_0 = s\}$ with $s_{t+1} \sim \mathcal{P}(s_t, a_t, \mathcal{L}_t)$. Here, $\{\mathcal{L}_t\}_{t=0}^{\infty}$ with $\mathcal{L}_t = \mathbb{P}_{s_t,a_t} \in \mathcal{P}(S \times A)$ has marginal distributions a_t for the population action and μ_t for the population state. According to this setting, also known as the general version of MFG, we can define the game's equilibrium according to the [157] as: given the game profile $G^* = (\{\mathcal{L}_t^*\}_{t=0}^{\infty}, \{\pi_t^*\}_{t=0}^{\infty})$, G^* represent a Nash equilibrium if

1. for any π_t and fixed $\{\mathcal{L}_t^*\}_{t=0}^{\infty}$ and initial state $s \in S$

$$\mathbb{E}\{\sum_{t=0}^{\infty} \gamma r_t^i(s_t, \pi_t^*(s_t, \mu_t(s_t)), \mathcal{L}_t^*)|s_0 = s\} \geq \mathbb{E}\{\sum_{t=0}^{\infty} \gamma r_t^i(s_t, \pi_t(s_t, \mu_t(s_t)), \mathcal{L}_t^*)|s_0 = s\}. \tag{4.4}$$

2. $\mathcal{P}_{s_t,a_t} = \mathcal{L}_t^*$ for all $t \geq 0$, where $st, at_{t=0}^{\infty}$ is the dynamics under the policy π^* starting from $s_0 \sim \mu_0^*$,with $at \sim \pi_t^*(st, \mu_t^*)$, $s_{t+1} \sim P(|s_t, a_t, \mathcal{L}_t^*)$, and$\mu_t^*$ being the population-state marginal of $\mathcal{L}_t^*$.

When $\{\mathcal{L}_t^*\}_{t=0}^{\infty}$ is fixed the condition 1 captures the optimal π^*. The condition 2 ensures the "consistency" of the solution that is, it guarantees that the state and action distribution flow of the single player does match the population state and action sequence $\mathcal{L}^*$. Further detail about MFGs can be found in [157].

4.1.2 Value and Q-value Functions, Optimal Policy, Q-Learning

In the section 4.1.1 we have introduced the expected long-run reward function that is used to manage the agent decision-making. By the way, it can be defined as follow:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + ... = \sum_{k=0}^{T} \gamma^k r_{t+k+1} \tag{4.5}$$

and called **"Return R_t"**. To find the optimal policy π^* as in 4.2, we have to define the **"Value-function V"**. This Value-function describes how it is beneficial for an agent to reach a given state s. This depends on the agent policy π of course. This value can be defined as:

$$V^{\pi}(s_t) = \mathbb{E}\{R_t|s_t = s\} = \mathbb{E}\{\sum_{k=0}^{T} \gamma^k r_{t+k+1}|s_t = s\} \tag{4.6}$$

Similarly, the **"Action-Value function Q"** represents the value of taking action a in state s under a policy π as:

$$Q^{\pi}(s_t, a_t) = \mathbb{E}\{R_t | s_t = s, a_t = a\} = \mathbb{E}\{\sum_{k=0}^{T} \gamma^k r_{t+k+1} | s_t = s, a_t = a\} \tag{4.7}$$

The presented problem in 4.2 can be rewritten according the definition of 4.7 or 4.6 as follow:

$$\pi^* = argmax_{\pi} Q^{\pi}(s, a) \tag{4.8}$$

$$\pi^* = argmax_{\pi} V^{\pi}(s) \tag{4.9}$$

To find the optimal policy π, we need to find the optimal values $max_{\pi} Q^{\pi}(s, a)$ or $max_{\pi} V^{\pi}(s)$. As described in [158] and [159],those optimal values are obtained by solving the **Bellman equation**:

$$Q^*(s_t, a_t) = r(s_t, a_t) + \gamma \mathbb{E}\{max_{a_{t+1}} Q^*(s_{t+1}, a_{t+1})\} \tag{4.10}$$

$$V^*(s_t) = \gamma \mathbb{E}_{\pi}\{r(s_t, a_t) + max_{a_t} V^*(s_{t+1})\} \tag{4.11}$$

4.1.3 Value-based Techniques

The most used equation is the eq. 4.10 and it can be solved by using the Value-based technique. The first one we introduce is the **Q-learning**. It retrieves the optimal values of the eq. 4.8 iteratively using the defined **Bellman equation updating rule** as:

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha_t [r_t(s_t, a_t) + \gamma max_{\alpha_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \tag{4.12}$$

where α_t is the learning rate that defines how much the new information contributes to the value of the existing $Q(s_t, a_t)$. The main idea behind the updating-rule 4.12 relies on finding the time difference between the current $Q(s_t, a_t)$ and the predicted $r_t(s_t, a_t) + \gamma max_{\alpha_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)$[3]. This algorithm uses the previous updating rule to build all possible Q-values for each (s_t, a_t) pair. The termination state is reached when a certain number of iterations are executed or when all Q-values have converged. In this case, we have the optimal policy π^* determining the optimal action to take in each state such that $Q^{\pi^*}(s_t, a_t)$ is maximized for each state in the space set.

[3]More details about the Q-learning algorithm can be found in [158].

In other terms, the policy π^* says that in any state we take the action that will get the highest cumulative reward. Q-learning technique suffers from the curse of dimensionality, in the sense that, this technique can be applied only to problems with low dimensionality of both state and action spaces. This afterwards makes Q-learning-based approaches not scalable. To overcome this limitation, the **Deep Q Network (DQN)** technique has been developed. This technique inherits the benefits of classical Q-learning. The main idea is to replace the table in the Q-learning algorithms with a deep neural network trying to approximate the Q values. The approximating function, represented by the deep neural network, is $Q(s_t, a_t, |\Phi)$ where Φ is the vector containing the vector parameters of the DNN. The fig.4.5 shows an example of DQN-

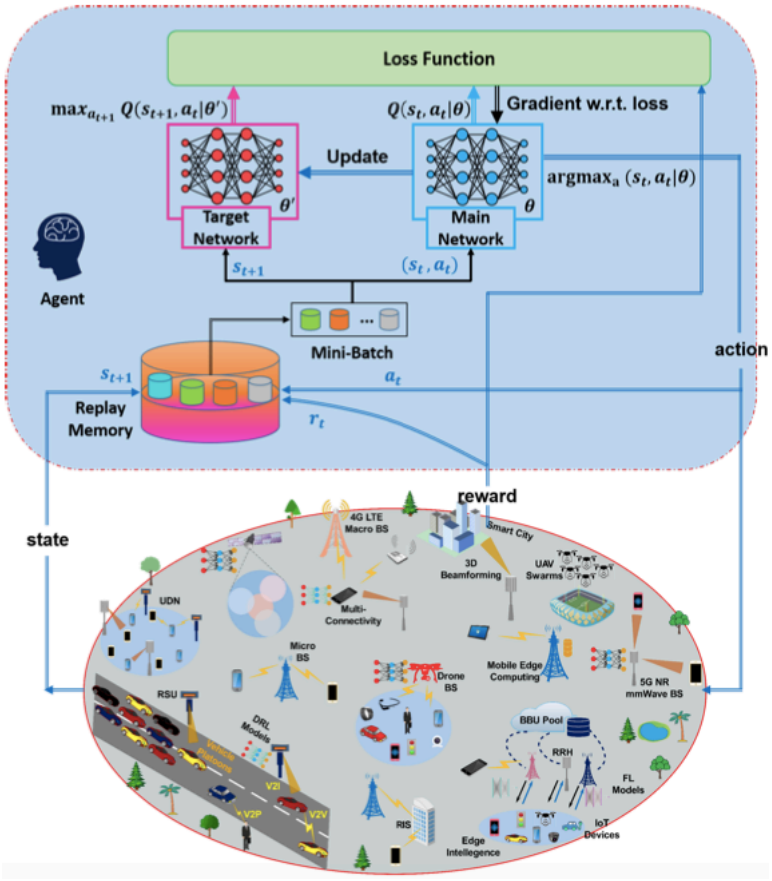


Figure 4.5: Deep Q network learning architecture example [154]

based architecture. Here the replay memory is denoted by D, and it is mainly used to break the correlation between the training samples, i.e., (s_t, a_t, r_t, s_{t+1}), by making them independently and identically distributed i.i.d. Those training samples are stored in the replay memory and the agent will select a mini-batch to train the neural network. To enhance the stability of the DQN model, the target Q network is used,

whose weights will be periodically updated to track those of the main Q network. The DQN approach is usually used to find the optimal policy π^*. This means that the Agent, by using the neural networks, will evaluate the derived optimal Q-value function from the iterative Bellman Equation:

$$Q(s_t, a_t) = r(s_t, a_t) + \gamma max_{a_{t+1}} Q(s_{t+1}, a_t) \tag{4.13}$$

and the DQN algorithm is then optimized by iteratively updating the weights Φ of its neural network to minimize the following Bellman loss function:

$$L(\Phi) = \mathbb{E}_{s_t, a_t, r_t, s_{t+1} \in D}[r(s_t, a_t) + \gamma max_{a_{t+1}} Q(s_{t+1}, a_t|\Phi') - Q(s_t, a_t|\Phi)]^2 \tag{4.14}$$

where Φ' is the vector containing the parameters of the Q target network. This kind of algorithm can be applied in each scenario with a discrete set of actions. However, there are limitations to its usage. This limitation emerges when the set of the action is continuous and it requires quantization. This, in turn, makes the DQN architecture quantization errors prone. Moreover, the DQN algorithm tends to overestimate the Q-values, which can degrade the training process and lead to sub-optimal policies. To enhance the DQN, **double DQN** has been developed. With the classical DQN, the overestimation results from the positive bias caused by the max operation employed in the Bellman equation. Specifically, the root cause is that the same training transitions are utilized in selecting and evaluating an action. In [160], the authors proposed a solution to this problem. They propose two Q-value functions. One is for selecting the best action and the other one is for evaluating the best action. The action selection is still based on the online training parameters Φ. Φ', instead, is used to evaluate the action. with this approach, the value of the policy is still estimated based on the current Q values and the training parameters are updated via switching between Φ and Φ'. In particular, according to [160], the evaluation of target Q value follows:

$$Q(s_t, a_t) = r(s_t, a_t) + \gamma Q(max_{a_{t+1}} Q(s_{t+1}, a_t|\Phi), \Phi') \tag{4.15}$$

and as Bellman loss function, double-DQN uses the following:

$$L(\Phi) = \mathbb{E}_{s_t, a_t, r_t, s_{t+1} \in D}[r(s_t, a_t) + Q(max_{a_{t+1}} Q(s_{t+1}, a_t|\Phi), \Phi') - Q(s_t, a_t|\Phi)]^2 \tag{4.16}$$

With these changes, double DQN solves the aforementioned problem of the $Q-value$ overestimation. However, it still suffers from some shortcomings. An alternative version providing an enhancement to the classical DQN and double DQN is represented by the presented approach named **Dueling DQN** in [161]. They proposed a new

point of view of the Q value function. They assess that this function can be expressed by the sum of two terms. One evaluates the being in a particular state, also known as state-value $V(s)$, and the second evaluates the importance of selecting a particular action given a state called $A(s,a)$. So, the Q value function is expressed as $Q(s,a) = V(s) + A(s,a)$, then they suggest having two main independent paths of fully-connected layers instead of having only a single path as the case in the basic DQN. These paths will estimate the terms $V(s)$ and $A(s,a)$ and then will be combined to provide $Q(s,a)$. Here the loss function is the same as in previous techniques, DQN and double DQN.

4.1.4 Policy-based Techniques

The policy-based techniques are part of the policy gradient family methods. They provide an alternative way to solve the MDP problems having high dimensionality and continuous action spaces. Recall that, the main object is the Q-value. If the Q-values are known, the optimal policy is defined by selecting actions that maximize the Q-values in each state. However, in environments with continuous action spaces, the Q function cannot be obtained as it is impossible to conduct a full search in a continuous action space to obtain the optimal action. Hence, the value-based approaches are inapplicable to problems characterized by their continuous action space, and the policy-based methods are applied instead. With the policy-based techniques, we avoid calculating Q values and directly obtain the optimal policy $\pi^* = \pi_{\Phi^*}$ that maximizes the agent's expected accumulated reward J defined as follows:

$$J(\Phi) = \mathbb{E}_{\pi_\Phi}[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)]^2 \tag{4.17}$$

The policy gradient approaches learn the optimal network training parameters Φ^* via performing gradient ascent on the function J.

$$\nabla_\Phi J(\Phi) = \mathbb{E}_{\pi_\Phi}[\sum_{t=0}^{T} \nabla log \pi_\Phi(a_t|s_t) Q^{\pi_\Phi}(s_t, a_t)] \tag{4.18}$$

In this formula, the function $Q^{\pi_\Phi}(s_t, a_t)$ is unknown. By the way, some algorithms have been developed to estimate this value and they are *REINFORCE Algorithm*, *Actor-Critic Algorithm*, *A3C Algorithm*, *Deep Deterministic Policy Gradient (DDPG) Algorithm*. The main idea of **The REINFORCE algorthm** is the increment of the certainty of good actions and reduction of the certainty of bad actions. Respect with Deep Q-learning approach, the REINFORCE algorithm does needn't a replay

buffer during training. Although this enhances its convergence speed, it needs much more interaction with the environment. Also, since it depends on the probabilities returned by the network, which incorporate uniform random agent behaviour, the REINFORCE algorithm performs the exploration process. Then, the target network isn't required in the REINFORCE algorithm as the Q values are obtained from the experiences in the environment. The disadvantage of the REINFORCE algorithm is that it suffers from high variance, meaning that any small shift in the return leads to a different policy. This limitation motivated the actor-critic algorithms. The main idea of **Actor-Critic Algorithm A2C** is the reduction of the variance that will enhance the convergence speed and stability of the policy gradient method. This approach utilises the accumulated discounted reward J to obtain the gradient of the policy J, which provides the direction that enhances the policy. The critical part of the algorithm reduces the variance of gradient estimates since it utilises various samples[4]. The selection of the best action in any state is based on $Q(s,a) = V(s) + A(s,a)$. A Bellman-based and trained deep neural network is used to estimate $V(s)$. Then, the estimated value is used to obtain the policy gradient and update the policy network such that the probabilities of actions with good advantage values are increased. Hence, the actor is the policy network $\pi(a|s)$ that takes actions by returning the probability distribution of actions, while the critic network evaluates the quality of the taken actions, $V(s)$. This algorithm is also called the advantage actor-critic method (A2C). The updated rules of the training parameters of the Actor and Critic neural networks follow the rule already shown in [158]. **The asynchronous advantage actor-critic algorithm A3C** is used to solve the high variance issue in gradients that results in non-optimal policies. The A3C algorithm conducts a parallel implementation of the actor-critic algorithm, where the actor and critic share the network layers. A global Neural network is trained to output action probabilities and an estimate of the *advantage function*. Several parallel actor learners are instantiated with copies of both the environment and training parameters of the global neural network. Each learner independently interacts with the environment and gathers training transitions to derive the gradients concerning its parameters. Learners then will propagate their gradients to the global neural network to update its weights. This mechanism ensures a periodic update of the global model with diverse transitions from each learner. The **Deep Deterministic Policy Gradient (DDPG) Algorithm**, represented by the fig.4.6, belongs to the actor-critic family, and it combines both Q-learning and policy gradients algorithms. It consists of actor and critic networks. The actor-network takes

[4]whereas the REINFORCE algorithm utilises only a single sample trajectory.

the state as its input, and it outputs the exact "deterministic" action, not probability distribution over actions as in the actor-critic algorithm. Whereas the critic is a Q value network that takes both the state and action as inputs and outputs the Q-value as a single output. As shown in [162], the DDPG is based on the DPG algorithm where both the policy and critic are deep neural networks. The DDPG algorithm creates a copy of both the actor and critic networks to compute the target values. The weights of these two target networks are then updated to slowly track the weight of the learned network to provide more stable training. Then the critic network is updated to minimize the Bellman loss function as in [162].

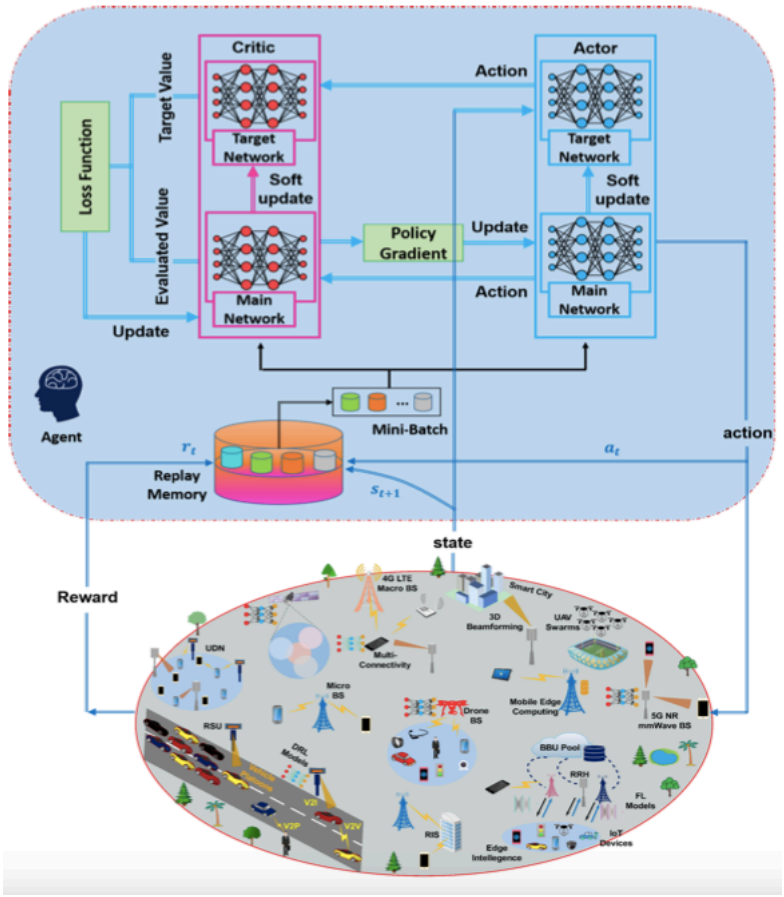


Figure 4.6: Deep Deterministic Policy Gradient Architecture provide in [154]

4.2 From Reinforcement to Deep Reinforcement Learning

In classical tabular RL, such as the Q-learning-based technique, usually, state and action spaces are small enough for the approximate value functions to be represented as

arrays or tables. In this condition, it is simple to find the optimal policy with specific methods, as in [158], inefficient way. However, these previous methods suffer from a difficult design issue when they come to real-world implementations. To overcome the problem of the tabular representation, we can design a parameterised value function with a weight vector. It is the case of a deep neural network that, at this point, can solve the problem without complete knowledge of the aforementioned state and action sets. When a deep neural network is used to implement the decision-making process of the agent, we are using a deep reinforcement learning methodology to solve a specific problem. DRL can complete complicated tasks with lower prior knowledge thanks to its ability to learn from scratch. So, the need of shifting from classical RL to DRL stems from the following considerations:

- **curse of dimensionality**. It refers to the fact that space and action sets cannot be stored following the tabular form. To overcome this issue, as in [163], function approximation is used to obtain features from models, value functions or policies and then attempts to generalise from them to construct an approximation of the entire function by deep neural networks.

- **Exploration-exploitation dilemma** In classical tabular-based RL, the uncertainty about the reward function and transition probabilities can be quantified as confidence intervals or posterior environment parameters. In DRL instead, an agent starts accumulating information about the environment, it has to navigate a trade-off between learning more about the environment (exploration) or pursuing the most promising strategy with the experience gathered (exploitation). Moreover, in DRL different settings are applied. One is that the agent explores only when the learning opportunities are valuable enough so that it can perform well without a separate training phase. Another one is that the agent follows a training policy during the first phase of interactions with the environment to accumulate training data and hence learn a test policy.

- **Convergence and stability**. For RL, only tables and linearly parameterised approximating functions can be used to guarantee convergence. When prior knowledge is not available to guide the selection of basis functions, a large number of basis functions must be defined to evenly cover the state-action space and this is impractical in high-dimensional problems. To address this problem, non-linear approximating functions, eg. Deep Neural networks, have been applied to obtain features of certain parts of states with replay buffers and target networks.

Given this consideration, for a practical system is preferable to use DRL-based approaches.

4.3 DRL-based Radio Resource management

The DRL-based approaches have been used to address different radio resource issues such as power allocation, and spectrum allocation. However, these approaches have been widely used to solve even other issues such as access control, rate control, and the joint use of these radio resources. In this section, we present a list of interesting works related to power and spectrum allocation[5][6].

4.3.1 DRL for Power Allocation

In this section, we present a list of interesting works related to the issue of power allocation that has been addressed with DRL-based approaches. The energy-efficient communication is achieved via efficient power allocation to ensure high QoS, better coverage, and enhanced data rate. Power allocation is involved in important network operations such as modulation and coding schemes, path loss compensation, interference management, etc. On the other hand, almost all modern user devices and IoT sensors are battery-powered with very limited battery capacity and charging capabilities. Hence, designing energy-efficient resource allocation schemes, protocols, and algorithms becomes fundamental in dynamic wireless network environments. The classical approaches are iterative and model-driven. Moreover, they are typically executed in a centralized fashion in which a network controller has full channel state information. In such a mechanism, BSs, wireless APs, and/or user devices require to wait until the centralized controller's iterations converge and send the outcome back over backhaul links. These approaches become impracticable in a large-scale system where is hard to obtain accurate and instantaneous channel state information. Hence, DRL techniques are used instead due to their superiority in obtaining optimal power allocation policies based on limited CSI. In [164] the authors propose a multi-agent Q-learning-based model to address the problem of downlink power control and rate adaptation in cellular systems. The agent is a network entity that resides in the cell, such as an eNB. The action space is discrete, corresponding to allocating downlink transmit power of the cell. The state space comprises four elements; cell power, average reference signal received power, average interference, and cell reward.

[5]To be coherent with the covered topic my PhD studies program

[6]About other issues that are not treated here, we invite the interested to read the work [154]

The reward function is continuous and defined based on the α-fair resource allocation utility function. It is defined by the performance indicator associated with all users in the radio cell. Their results show that the agent quickly learns the power control policy to provide significant energy savings and fairness across the network users. In [165], the authors propose single and multi-agent actor-critic DRL methods to address the problem of downlink sum-rate maximization through power allocation in multi-cell, multi-user cellular networks. In their model, the agents are the BSs. Their state space is continuous and comprises network CSI and the transmit power allocation by previous BSs. The action space is continuous too and corresponds to the power allocation, while the reward function is the cellular network sum SE. Experimental results demonstrate that their DRL-based method can achieve higher SE than conventional optimisation algorithms. The authors of [166] address the problem of throughput maximisation in CRANs via power allocation in vitalised 5G networks. They propose a multi-agent DQN-based DRL algorithm to solve the problem in which the agents are each linked between RRH and the user. The action space corresponding to the transmit power is discrete, The state space representing the current partial CSI and the respective power set is continuous. The reward of each slice is discrete, defined as a function of the sum of its tenants' rates. Given this model, the results of their simulation show that their proposed scheme achieves a higher sum rate compared to greedy search-based power allocation approaches. The authors in [167] design a multi-agent DQN and DDPG-based DRL framework to address the problem of power control in a Heterogenous network. In this model, the APs are the agents and they are equipped with a local DNN. The state space of each local DNN is continuous and represents the local state information, while the local action space is continuous and represents the transmit power. The method they propose is called multiple-actor-shared-critic (MASC). Its main idea is the following: we have a collection of multiple and identical actor DNNs and only a shared critic DNN. Historical global information is provided to the critic DNN, and the output of the critic DNN will evaluate whether the output power of each actor DNN is optimal or not from a global view. The reward function is continuous and is based on the data rate between each AP and its associated user. Simulation results show that their proposed algorithm outperforms better respect with their competitors in terms of both convergence rate and computational complexity. In [168] proposes a distributive multi-agent DDPG-based DRL algorithm to address the problem of sum-rate maximization via continuous power control in wireless mobile networks. In this model, the transmitter is considered as the agent and the agent's state is a combination of features such as the

local information, interfering neighbours, and interfered neighbours groups. The action is expressed by the choice of the transmit power level, while the reward function is based on the sum rate of the sum-rate maximisation problem. Their results show that the method gives better performance than the conventional approaches. In [169] the authors propose a data-driven approach based on a multi-agent DQN algorithm to address the downlink power control in dense 5G cellular networks. Here, the agents are the BSs, whose state space is continuous. It is composed of path gain, SINR, Dl rate, and downlink power. The action space is discrete and it is represented by the Dl power, while the reward is a function of the network-wide harmonic mean of throughput. Their simulation results show that their approach performs better than the fixed power allocation approach. In particular, this method can improve data rates at the cell edge while ensuring reduced transmitted power.

4.3.2 DRL for Spectrum Allocation

In this section, we present a list of interesting works related to the issue of spectrum allocation that has been addressed with DRL-based approaches. In modern wireless networks, a massive number of user devices may request to simultaneously access the wireless channel. This may imply overload and congest the channel causing communication failure and unreliable QoS. Hence, efficient communication schemes and protocols must be developed to address this issue. These schemes may include various access scheduling and prioritization techniques. Moreover, modern wireless networks require dynamic load balancing and access management methods to support the massive capacity and connectivity requirements of future wireless networks while utilizing their radio resources efficiently. Anyway, the major part of the optimization-based models which you can define to meet those requirements, belong to the set of combinatorial and no-convex problems requiring perfect information to obtain optimal solutions. The acquisition of this information is impracticable in real systems and this forced the network engineer to use other kinds of methodologies. In this context, DRL-based techniques have attracted considerable interest recently due to their robustness in making optimal decisions in dynamic and stochastic environments such as wireless networks. In [169] the authors proposed a distributed multi-agent DQN. In their approach, the agents are each user. The action space is discrete and includes the transmit power change for each channel. The state space is discrete and corresponds to the transmit power on wireless channels. The reward is a continuous function based on the spectrum efficiency and the penalty related to the interference with

primary users. Experimental results show that the proposed model with network-based DQN achieves a higher reward with both the achievable data rate and primary user protections. The authors in [170] presented a DRL-based model to address the problem of dynamic spectrum allocation in distributed wireless networks. They design a channel access scheme to maximize channel throughput according to fair channel access. The agents are the users. The action space is discrete and includes the action "0" indicating if the user does not attempt to transmit packets during the current time slot, and the action 1 indicating if it has attempted to transmit. The state space is also discrete and consists of four main elements that are

- each user action was taken on the current time slot;
- channel capacity (which could be negative, positive, or zero);
- a binary acknowledgement signal showing if the user transmits successfully or not;
- a parameter that enables each user to estimate other users' situations.

The reward is a discrete binary function that takes the value of "1" if the user transmits successfully; otherwise, it is "0" meaning that the user transmitted with collision. Their results show that the proposed DRL-based model can maximize the total throughput while trying to make fair resource allocation among users. In [171] the author proposes a multi-agent deep recurrent Q-network-based model to address the problem of dynamic spectrum allocation in dynamic heterogeneous environments with partial observations. The considered scenario includes users sharing multiple independent radio channels. The users are the agents. The action space is discrete and it includes the chances of transmitting in a particular band or waiting during the next time slot. The state-space set is discrete and it includes the information indicating whether the channels are occupied, idle, or unknown. The reward function is discrete. It returns two per-channel values that are 100 for successful transmission and -50 for collision. With their results, the authors show that the proposed model handles various dynamic communication environments, and its performance outperforms the "myopic" conventional methods. Moreover, the performance is very close to classical and optimisation-model-based approaches. In [172], the authors present a multi-agent DRL network model combining recurrent neural networks and DQN that addresses the problem of finding the optimal dynamic spectrum allocation policy that jointly maximizes channel utilization and minimizes the packet loss rate. In the

model, each user acts as an agent which has a discrete action space including the choice of a channel for transmission at time slot t. The state space is discrete and it is composed of three components that are

1. a binary transmission condition η, which is 1 if the transmission is successful and 0 otherwise;
2. the channel selection action
3. the channel status indicator after each dynamic access process.

The reward is a discrete binary function which takes a positive value if $\eta = 1$, otherwise it takes the value of 0. Their results show that their model

1. reduces the packet loss rate from 0.8 to around 0.1;
2. outperforms Slotted-Aloha and DQN in terms of reducing collision probability and channel idle probability by about 60%;
3. enhances the transmission success rate by around 20%.

The last work that we present is [173]. Here, the authors addressed the problem of dynamic spectrum allocation in cognitive radio networks. To solve this problem, they proposed an uncoordinated and distributed multi-agent DQN model. In this model, the cognitive radios are the agents. They have a discrete action space including the possible transmit powers. Also, the state space is discrete and it includes a state indicating whether the limits for dynamic spectrum allocation are being met or not. This value depends on the relative throughput change at all the primary network links. Moreover, they use a discrete reward function based on the throughput of the links and the environment's state. Their experimental results show that their model finds policies that yield performance within 3% of an exhaustive search solution, and it finds the optimal policy in nearly 70% of cases.

4.4 Example of RL-based-modelling applied to the presented problems

In this section, we just provide the main idea of how to apply the RL paradigm to the presented problems. In particular, we provide a way to describe the Markov decision process that can be used to model the presented problems. The main purpose of this work is to demonstrate that the RL approaches can perform well in this context(in

the sense that they solve the problem first) and then, by comparison, they can even perform better concerning the classical approaches.

4.4.1 Dynamic selection of numerology and bandwidth part for effective ORAN slicing

The first step for applying the RL paradigms is the definition of the basic elements composing the Markov decision process which are: the States Set, Actions Set and Reward function. Given the nature of the system is impossible to define a transition probability function. Given the considered system cellular, its state S corresponds to already defined. So, $S = \{S_{SA}, S_{PRA}\}$. About the Actions Set, we can consider the one including the change of the BWP selecting algorithm. So, this set can be defined as follow: $A = \{a1, a2\}$ with $a1$ = "Take over SA algorithms with PRA" and $a2$ = "Take over PRA algorithms with SA". About the reward function, we can use the α-fair function as in [174]. This function can be defined as follow:

$$R(x) = \frac{1}{1-\alpha} \sum_{x_i \in X} w_i (h(x_i)^{1-\alpha} - 1) \tag{4.19}$$

with $\alpha \in (0, \infty)$ as a scalar coefficient, $h : X \to \mathbb{R}$ as a transfer function. The X is the set of all radio measurements or performance indicators associated with the radio-observed environment and used for the definition of the performance measurement. The $x = [x_1, \ldots, x_{|X|}]$ is the vector of performance indicators belonging to the set X. So, in this work, a possible definition of the performance indicators' set is the following $X = [L_{UP}, T_h]$. The w_i value is the weight associated with each performance indicator. Hence, h has to be designed to account for all the features in X. When this function refers to the single user i each reward $R_t(x)$ can be approximated to a different values, according to the values of α and weights w_i as described in [164]. This enables the agent node to optimize a different performance metric that can either be associated with individual user devices or the whole cellular radio network. According to the RL theory, the variable is used to define the state. to solve the problem we declared in the dedicated section of the work, we need to use an algorithm that doesn't need to know all the information about the observed environment. This means that they needn't store all the possible combinations of state-action in a tabular. The most promising methodology that can be used is DQN or DDQN.

4.4.2 Effective Multiplexing of URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity

Following the reason of the previous section, even for this work, we have to convert the system model into something that can be managed according to the RL framework. So, we need to describe the observed scenario following the main idea of the Markov Decision Process. Therefore, we need

- State set S. The state of the observed system can be defined as follow: $S = \{S_{Embb}, S_{Urllc}\}$ with $S_{Embb} = \{Th_1, Th_2, \ldots, Th_{|K_E|}\}$ and $S_{Urllc} = \{Th_1, Th_2, \ldots, Th_{|K_U|}\}$ representing the value of the observed throughout in the system on per-service based.

- Action set $A = \{a1, a2, a3, a4\}$ representing all the possible action which the agent can execute in the observed system. In particular

  - $a1$ = "Change pilot procedure assignment from CGPA to CRPA"
  - $a2$ = "Change pilot procedure assignment from CGPA to CRPA"
  - $a3$ = "Set Equal Power Allocation Scheme $v = 0$"
  - $a4$ = "Set Channel gain based Power Allocation Scheme $v = 1$"

- Even here we can use the α-fairness function. Our transformation function h is related to the **outage probability** metrics that is the function of the throughput of the user. These performance metrics are associated with the event describing a situation where the **current operational condition of the network doesn't sustain the transmission of packets of fixed size(b_u for URLLC and b_e for EMBB)**. This means that the maximization of the reward is constrained by the reference values, which are R_u^{th} and R_e^{th}, representing the requirements of both considered services in the system. According to the values of α it is possible to design agents with different levels of behaviours related to the fairness of the resource allocation.

Even in this case, since the variables describe the environment the agent has to interact with, we cannot use tabular-based reinforcement learning approaches. We have to equip our agent with a Deep neural network to approximate the values of each pair action state that will be used to find the optimal policy to solve the considered task related to resource management. So, the most promising approaches, that have been selected to be the decision-making core of the agent are DQN and DDQN.

Chapter 5

Summary

In this chapter, we present the theoretical elements of the new promising Artificial intelligence-based approach, that is reinforcement learning, leading the conversion of 5G into an automatic ecosystem never seen before. The potential of reinforcement learning has been further increased when it's been combined with deep learning. This is the case of Deep reinforcement learning. Given the width of the topic we limited the argumentation of deep reinforcement learning to the power and spectrum allocation. Accordingly, we reviewed the most used algorithms and mechanisms. Then we proposed the eventual application of the presented approach to our previous works. The research activity is still on going.

Chapter 6

Concluding remarks and future directions

This broad study on 5G radio access networks has evidenced a few techniques that can be used to enable flexibility in resource management with a consequent improvement in performance. The main approaches, introducing the required flexibility for the next generation of the radio access networks, are softwarization expressing mainly itself with the application of new software network paradigms eg. NFV and SDN and lastly with the NS and their alternative RL-based variation, and the adoption of new radio access network topology such the cell-free massive MIMO. For network-paradigms-based softwarization, we provide the presented work in sec 2.4 showing that the benefits associated with the extension of the SDN into the RAN are remarkable and can justify the softwarization of the RAN segment. Instead, for the flexible radio topology-based- approach, we presented the work in sec 3.3 where we show that even this methodology produces remarkable benefits justifying the deployment of this cell-free based scheme in the next generation of the radio access network. Moreover, the acquired expertise allows me to define specific future directions for the presented work. One of them is the re-arrangement of the studied problems according to the Reinforcement Learning-based fashion way. Other future directions include the design of smarter resources allocating mechanisms considering additional information concerning the ones that we have used here. Of course, this will lead to comparison studies between these competing algorithms that will be based on specific key performance indicators. Contemporary to this, this broad study allowed me to brush up on capabilities already acquired during my university career and acquire new ones which will be useful for the next future.

This thesis is dedicated in loving memory of Bucci Ernesto.
As I promised to both of us,
Dad, I've finally done it.

6.1 Acronyms

QoS Quality of Service 6

Ul Up-link ii

Dl down-link ii

CAPEX Capital Expenditures 7

OPEX Operational Expenditures 7

RAN Radio Access Network 6

CN Core Network 20

4Gn 4G Network 6

5Gn 5G Network 6

5G MCS Fifth Generation of Mobile Communication System 6

CRAN Cloud-RAN 6

H-CRAN Heterogeneous Cloud-RAN 6

V-CRAN Virtualized Cloud-RAN 6

List of Publications

Papers:

C1 U. Bucci, A. Marotta, K. Kondepu, C. Antonelli, L. Valcarenghi and D. Cassioli, "Experimental Demonstration of Flexible TDM-PON Fronthauling," 2018 Asia Communications and Photonics Conference (ACP), 2018, pp. 1-3, doi: 10.1109/ACP.2018.8596093.

J1 Calcaterra Claudio, Carmenini Alessio, Marotta Andrea, BuccI Ubaldo, Cassioli Dajana. (2020). MaxHadoop: An Efficient Scalable Emulation Tool to Test SDN Protocols in Emulated Hadoop Environments. Journal of Network and Systems Management, Springer. 28. 10.1007/s10922-020-09552-x.

J2 U. Bucci, A. Marotta and D. Cassioli, "Dynamic Bandwidth-Part Allocation for 5G services," Journal of Network and Systems Management, Springer (SUBMITTED TO Journal of Network and Systems Management).

J3 U. Bucci, D. Cassioli and A. Marotta, "Multiplexing URLLC and eMBB Traffic by Cell-Free Massive MIMO Spatial Diversity" IEEE Transactions on Network and Service Management (SUBMITTED TO IEEE Transactions on Network and Service Management).

References

[1] Y. Lin, L. Shao, Z. Zhu, Q. Wang, and R. K. Sabhikhi, "Wireless network cloud: Architecture and system requirements," *IBM J. Res. Dev.*, vol. 54, p. 4, 2010.

[2] "C-ran the road towards green ran,"

[3] G. Kardaras and C. Lanzani, "Advanced multimode radio for wireless amp; mobile broadband communication," in *2009 European Wireless Technology Conference*, pp. 132–135, 2009.

[4] M. A. Habibi, M. Nasimi, B. Han, and H. D. Schotten, "A comprehensive survey of ran architectures toward 5g mobile communication system," *IEEE Access*, vol. 7, pp. 70371–70421, 2019.

[5] M. Fiorani, S. Tombaz, J. Martensson, B. Skubic, L. Wosinska, and P. Monti, "Modeling energy performance of c-ran with optical transport in 5g network scenarios," *Journal of Optical Communications and Networking*, vol. 8, no. 11, pp. B21–B34, 2016.

[6] Z. Tan, C. Yang, and Z. Wang, "Energy evaluation for cloud ran employing tdm-pon as front-haul based on a new network traffic modeling," *Journalof Lightwave Technology*, vol. 35, no. 13, pp. 2669–2677, 2017.

[7] V. N. Ha, L. B. Le, and N.-D. Dào, "Energy-efficient coordinated transmission for cloud-rans: Algorithm design and trade-off," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2014.

[8] J. You, Z. Zhong, G. Wang, and B. Ai, "Security and reliability performance analysis for cloud radio access networks with channel estimation errors," *IEEE Access*, vol. 2, pp. 1348–1358, 2014.

[9] D. B. Rawat, S. Shetty, and K. Raza, "Secure radio resource management in cloud computing based cognitive radio networks," in *2012 41st International Conference on Parallel Processing Workshops*, pp. 288–295, 2012.

[10] B. Niu, Y. Zhou, H. Shah-Mansouri, and V. W. S. Wong, "A dynamic resource sharing mechanism for cloud radio access networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8325–8338, 2016.

[11] V. Suryaprakash, P. Rost, and G. Fettweis, "Are heterogeneous cloud-based radio access networks cost effective?," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2239–2251, 2015.

[12] L. Chen, L. Liu, X. Fan, J. Li, C. Wang, G. Pan, J. Jakubowicz, and T.-M.-T. Nguyen, "Complementary base station clustering for cost-effective and energy-efficient cloud-ran," in *2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, pp. 1–7, 2017.

[13] M. De Andrade, M. Tornatore, A. Pattavina, A. Hamidian, and K. Grobe, "Cost models for baseband unit (bbu) hotelling: From local to cloud," in *2015 IEEE 4th International Conference on Cloud Networking (CloudNet)*, pp. 201–204, 2015.

[14] O. Arouk, T. Turletti, N. Nikaein, and K. Obraczka, "Cost optimization of cloud-ran planning and provisioning for 5g networks," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2018.

[15] H. Niu, C. Li, A. Papathanassiou, and G. Wu, "Ran architecture options and performance for 5g network evolution," in *2014 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 294–298, 2014.

[16] M. Fathy, B. Mokhtar, M. A. Abdou, and M. R. M. Rizk, "Extended study towards performance improvement of cloud-ran," in *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 1061–1066, 2017.

[17] F. A. Khan, H. He, J. Xue, and T. Ratnarajah, "Performance analysis of cloud radio access networks with distributed multiple antenna remote radio heads," *IEEE Transactions on Signal Processing*, vol. 63, no. 18, pp. 4784–4799, 2015.

[18] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, "Cloud radio access network (c-ran): a primer," *IEEE Network*, vol. 29, no. 1, pp. 35–41, 2015.

[19] U. Karneyenka, K. Mohta, and M. Moh, "Location and mobility aware resource management for 5g cloud radio access networks," in *2017 International Conference on High Performance Computing Simulation (HPCS)*, pp. 168–175, 2017.

[20] L. Liu, F. Yang, R. Wang, Z. Shi, A. Stidwell, and D. Gu, "Analysis of handover performance improvement in cloud-ran architecture," in *7th International Conference on Communications and Networking in China*, pp. 850–855, 2012.

[21] M. Peng, Y. Li, J. Jiang, J. Li, and C. Wang, "Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 126–135, 2014.

[22] Y. Li, T. Jiang, K. Luo, and S. Mao, "Green heterogeneous cloud radio access networks: Potential techniques, performance trade-offs, and challenges," *IEEE Communications Magazine*, vol. 55, no. 11, pp. 33–39, 2017.

[23] J. Li, M. Peng, Y. Yu, and Z. Ding, "Energy-efficient joint congestion control and resource optimization in heterogeneous cloud radio access networks," 2016.

[24] R. S. Alhumaima, M. Khan, and H. S. Al-Raweshidy, "Power model for heterogeneous cloud radio access networks," in *2015 IEEE International Conference on Data Science and Data Intensive Systems*, pp. 260–267, 2015.

[25] Q.-T. Vien, T. Le, B. Barn, and P. Van Ca, "Optimising energy efficiency of non-orthogonal multiple access for wireless backhaul in heterogeneous cloud radio access network," *IET Communications*, vol. 10, pp. 2516 – 2524, 12 2016.

[26] H. Dahrouj, A. Douik, O. Dhifallah, T. Y. Al-Naffouri, and M.-S. Alouini, "Resource allocation in heterogeneous cloud radio access networks: advances and challenges," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 66–73, 2015.

[27] M. A. Marotta, N. Kaminski, I. Gomez-Miguelez, L. Z. Granville, J. Rochol, L. DaSilva, and C. B. Both, "Resource sharing in heterogeneous cloud radio access networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 74–82, 2015.

[28] B. Zhang, X. Mao, J.-L. Yu, and Z. Han, "Resource allocation for 5g heterogeneous cloud radio access networks with d2d communication: A matching and coalition approach," *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 5883–5894, 2018.

[29] S.-Y. Lien, S.-M. Cheng, K.-C. Chen, and D. I. Kim, "Resource-optimal licensed-assisted access in heterogeneous cloud radio access networks with heterogeneous carrier communications," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9915–9930, 2016.

[30] X. Wang, C. Cavdar, L. Wang, M. Tornatore, H. S. Chung, H. H. Lee, S. M. Park, and B. Mukherjee, "Virtualized cloud radio access network for 5g transport," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 202–209, 2017.

[31] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," *IEEE Communications Magazine*, vol. 51, no. 7, pp. 27–35, 2013.

[32] ITU-T, "5g wireless fronthaul requirements in a passive optical network context," *ITU-T G-series Recommendations – Supplement 66*, 2020.

[33] S. Costanzo, D. Xenakis, N. Passas, and L. Merakos, "Openb: A framework for virtualizing base stations in lte networks," 06 2014.

[34] H. Li, M. Dong, and K. Ota, "Radio access network virtualization for the social internet of things," *IEEE Cloud Computing*, vol. 2, no. 6, pp. 42–50, 2015.

[35] G. Dandachi, T. Chahed, S. E. Elayoubi, N. C. Taher, and Z. Fawal, "Joint allocation strategies for radio and processing resources in virtual radio access networks (v-ran)," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2017.

[36] L. Tian, Y. Zhou, Y. Wang, J. Yang, Q. Sun, J. Yuan, and B. Yang, "Evaluation methodology for virtual base station platforms in radio access networks," *IEEE Access*, vol. 6, pp. 49366–49374, 2018.

[37] G. Tseliou, F. Adelantado, and C. Verikoukis, "Scalable ran virtualization in multitenant lte-a heterogeneous networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6651–6664, 2016.

[38] I. Al-Samman, M. Artuso, H. Christiansen, A. Doufexi, and M. Beach, "A framework for resources allocation in virtualised c-ran," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–7, 2016.

[39] D. Mishra, H. Gupta, B. R. Tamma, and A. A. Franklin, "Kora: A framework for dynamic consolidation amp; relocation of control units in virtualized 5g ran," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–7, 2018.

[40] W. Al-Zubaedi and H. S. Al-Raweshidy, "A parameterized and optimized bbu pool virtualization power model for c-ran architecture," in *IEEE EUROCON 2017 -17th International Conference on Smart Technologies*, pp. 38–43, 2017.

[41] P. Rost, I. Berberana, A. Maeder, H. Paul, V. Suryaprakash, M. Valenti, D. Wübben, A. Dekorsy, and G. Fettweis, "Benefits and challenges of virtualization in 5g radio access networks," *IEEE Communications Magazine*, vol. 53, no. 12, pp. 75–82, 2015.

[42] M. Mukherjee, L. Shu, and D. Wang, "Survey of fog computing: Fundamental, network applications, and research challenges," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 1826–1857, 2018.

[43] R. K. Naha, S. Garg, D. Georgakopoulos, P. P. Jayaraman, L. Gao, Y. Xiang, and R. Ranjan, "Fog computing: Survey of trends, architectures, requirements, and research directions," *IEEE Access*, vol. 6, pp. 47980–48009, 2018.

[44] S. Jia, Y. Ai, Z. Zhao, M. Peng, and C. Hu, "Hierarchical content caching in fog radio access networks: ergodic rate and transmit latency," *China Communications*, vol. 13, no. 12, pp. 1–14, 2016.

[45] H. Xiang, M. Peng, Y. Cheng, and H.-H. Chen, "Joint mode selection and resource allocation for downlink fog radio access networks supported d2d," in *2015 11th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QSHINE)*, pp. 177–182, 2015.

[46] S.-H. Park, O. Simeone, and S. Shamai, "Joint cloud and edge processing for latency minimization in fog radio access networks," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, 2016.

[47] R. Tandon and O. Simeone, "Cloud-aided wireless networks with edge caching: Fundamental latency trade-offs in fog radio access networks," in *2016 IEEE International Symposium on Information Theory (ISIT)*, pp. 2029–2033, 2016.

[48] Y. Qiu, H. Zhang, K. Long, Y. Huang, X. Song, and V. C. M. Leung, "Energy-efficient power allocation with interference mitigation in mmwave-based fog radio access networks," *IEEE Wireless Communications*, vol. 25, no. 4, pp. 25–31, 2018.

[49] K. Liang, L. Zhao, X. Zhao, Y. Wang, and S. Ou, "Joint resource allocation and coordinated computation offloading for fog radio access networks," *China Communications*, vol. 13, no. Supplement2, pp. 131–139, 2016.

[50] H. Zhang, Y. Qiu, K. Long, G. K. Karagiannidis, X. Wang, and A. Nallanathan, "Resource allocation in noma-based fog radio access networks," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 110–115, 2018.

[51] M. Peng and K. Zhang, "Recent advances in fog radio access networks: Performance analysis and radio resource allocation," *IEEE Access*, vol. 4, pp. 5003–5009, 2016.

[52] T.-C. Chiu, W.-H. Chung, A.-C. Pang, Y.-J. Yu, and P.-H. Yen, "Ultra-low latency service provision in 5g fog-radio access networks," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2016.

[53] J. Sasiain, A. Sanz, J. Astorga, and E. Jacob, "Towards flexible integration of 5g and iiot technologies in industry 4.0: A practical use case," *Applied Sciences*, vol. 10, no. 21, 2020.

[54] Z. Ma, Z. Zhang, Z. Ding, P. Fan, and H. Li, "Key techniques for 5g wireless communications: Network architecture, physical layer, and mac layer perspectives," *Science China Information Sciences*, vol. 58, pp. 1–20, 04 2015.

[55] I. Akyildiz, p. Wang, and S.-C. Lin, "Softair: A software defined networking architecture for 5g wireless systems," *Computer Networks*, vol. 85, 05 2015.

[56] G. Sun, F. Liu, J. Lai, and G. Liu, "Software defined wireless network architecture for the next generation mobile communication: Proposal and initial prototype," *Journal of Communications*, vol. 9, 12 2014.

[57] B. A. A. Nunes, M. Mendonca, X.-N. Nguyen, K. Obraczka, and T. Turletti, "A survey of software-defined networking: Past, present, and future of programmable networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1617–1634, 2014.

[58] "Network function virtualisation-introductory white paper," *SDN and Openflow World Congr*, 2012.

[59] J. Matias, J. Garay, N. Toledo, J. Unzilla, and E. Jacob, "Toward an sdn-enabled nfv architecture," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 187–193, 2015.

[60] W. Ding, W. Qi, J. Wang, and B. Chen, "Openscaas: an open service chain as a service platform toward the integration of sdn and nfv," *IEEE Network*, vol. 29, no. 3, pp. 30–35, 2015.

[61] L. García Villalba, L. Valdivieso, D. Lopez, and L. Barona, "Trends on virtualisation with software defined networking and network function virtualisation," *IET Networks*, vol. 4, 08 2015.

[62] Y. Li and M. Chen, "Software-defined network function virtualization: A survey," *IEEE Access*, vol. 3, pp. 2542–2553, 2015.

[63] J. d. J. Gil Herrera and J. F. Botero Vega, "Network functions virtualization: A survey," *IEEE Latin America Transactions*, vol. 14, no. 2, pp. 983–997, 2016.

[64] ETSI, "5g;ng-ran; architecture description," *TS*, vol. V15.2.0, 2018.

[65] M. A. Habibi, B. Han, and H. Schotten, "Network slicing in 5g mobile communication architecture, profit modeling, and challenges," 07 2017.

[66] B. Han, S. Tayade, and H. D. Schotten, "Modeling profit of sliced 5g networks for advanced network resource management and slice implementation," in *2017 IEEE Symposium on Computers and Communications (ISCC)*, pp. 576–581, 2017.

[67] D. Bega, M. Gramaglia, A. Banchs, V. Sciancalepore, K. Samdanis, and X. Costa-Perez, "Optimising 5g infrastructure markets: The business of network slicing," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pp. 1–9, 2017.

[68] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira, "Network slicing for 5g with sdn/nfv: Concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, 2017.

[69] M. A. Albreem, M. Juntti, and S. Shahabuddin, "Massive mimo detection techniques: A survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3109–3132, 2019.

[70] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *CoRR*, vol. abs/1710.00395, 2017.

[71] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172–179, 2013.

[72] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive mimo linear precoding: A survey," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3920–3931, 2018.

[73] Z. Liu and L. Dai, "A comparative study of downlink mimo cellular networks with co-located and distributed base-station antennas," *IEEE Transactions on Wireless Communications*, vol. 13, no. 11, pp. 6259–6274, 2014.

[74] J. Wang and L. Dai, "Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3046–3058, 2015.

[75] S. Elhoushy, M. Ibrahim, and W. Hamouda, "Cell-free massive mimo: A survey," *IEEE Communications Surveys Tutorials*, vol. 24, no. 1, pp. 492–523, 2022.

[76] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo: Uniformly great service for everyone," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 201–205, 2015.

[77] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[78] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[79] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive mimo systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 7, pp. 4445–4459, 2017.

[80] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *CoRR*, vol. abs/1903.10611, 2019.

[81] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *CoRR*, vol. abs/2108.02541, 2021.

[82] ETSI, "Lte; evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran); overall description; stage 2," *TS*, vol. V11.6.0, 2013.

[83] Z. E. Ankarali, B. Peköz, and H. Arslan, "Flexible radio access beyond 5g: A future projection on waveform, numerology, and frame design principles," *IEEE Access*, vol. 5, pp. 18295–18309, 2017.

[84] A. Sahin and H. Arslan, "Multi-user aware frame structure for ofdma based system," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1–5, 2012.

[85] M. Iwabuchi, A. Benjebbour, Y. Kishiyama, D. Wu, T. Tian, L. Gu, Y. Cui, and T. Kashima, "5g field experimental trial on frequency domain multiplexing of mixed numerology," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2017.

[86] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. Da Silva, "Waveform and numerology to support 5g services and requirements," *IEEE Communications Magazine*, vol. 54, no. 11, pp. 90–98, 2016.

[87] A. Ijaz, L. Zhang, M. Grau, A. Mohamed, S. Vural, A. U. Quddus, M. A. Imran, C. H. Foh, and R. Tafazolli, "Enabling massive iot in 5g and beyond systems: Phy radio frame design considerations," *IEEE Access*, vol. 4, pp. 3322–3339, 2016.

[88] "Share technote." `https://www.sharetechnote.com/html/5G/5G_FrameStructure.html`. Accessed: 2022-05-07.

[89] S. Coltellacci, D. Franci, E. Grillo, S. Pavoncello, S. d'Elia, N. Pasquino, M. Vaccarono, S. Adda, and R. Suman, "A methodology to characterize power control systems for limiting exposure to electromagnetic fields generated by massive mimo antennas," *IEEE Access*, vol. 8, pp. 171956 – 171967, 09 2020.

[90] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. USA: Academic Press, Inc., 1st ed., 2018.

[91] "Bandwidth part adaptation 5 g nr user experience & power consumption enhancements,"

[92] X. Lin, D. Yu, and H. Wiemann, "A primer on bandwidth parts in 5g new radio," *ArXiv*, vol. abs/2004.00761, 2020.

[93] K. Takeda, H. Xu, T. Kim, K. Schober, and X. Lin, "Understanding the heart of the 5g air interface: An overview of physical downlink control channel for 5g new radio," *IEEE Communications Standards Magazine*, vol. 4, no. 3, pp. 22–29, 2020.

[94] 3GPP, "5g; nr; physical layer procedures for control," *TS*, vol. V15.3.0, 2018.

[95] "Howltestuffworks." `http://howltestuffworks.blogspot.com/2019/11/5g-nr-bandwidth-part.html`. Accessed: 2022-05-15.

[96] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: Software defined radio access network," in *Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking*, HotSDN '13, (New York, NY, USA), p. 25–30, Association for Computing Machinery, 2013.

[97] M. Y. Arslan, K. Sundaresan, and S. Rangarajan, "Software-defined networking in cellular radio access networks: potential and challenges," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 150–156, 2015.

[98] X. Foukas, N. Nikaein, M. M. Kassem, M. K. Marina, and K. Kontovasilis, "Flexran: A flexible and programmable platform for software-defined radio access networks," in *Proceedings of the 12th International on Conference on Emerging Networking EXperiments and Technologies*, CoNEXT '16, (New York, NY, USA), p. 427–441, Association for Computing Machinery, 2016.

[99] R. Riggio, M. K. Marina, J. Schulz-Zander, S. Kuklinski, and T. Rasheed, "Programming abstractions for software-defined wireless networks," *IEEE Transactions on Network and Service Management*, vol. 12, no. 2, pp. 146–162, 2015.

[100] N. Nikaein, M. K. Marina, S. Manickam, A. Dawson, R. Knopp, and C. Bonnet, "Openairinterface: A flexible platform for 5g research," *SIGCOMM Comput. Commun. Rev.*, vol. 44, p. 33–38, oct 2014.

[101] I. Gomez-Miguelez, A. Garcia-Saavedra, P. D. Sutton, P. Serrano, C. Cano, and D. J. Leith, "srslte: An open-source platform for lte evolution and experimentation," 2016.

[102] ONF, "Sd-ran: Onf's software-defined ran platform consistent with the o-ran architecture," *White Paper*, 2020.

[103] 3GPP, "Nr; physical channels and modulation," *TS*, vol. V 38.211, 2021.

[104] N. Patriciello, S. Lagén, B. Bojovic, and L. Giupponi, "An E2E simulator for 5g NR networks," *CoRR*, vol. abs/1911.05534, 2019.

[105] ETSI, "Nr; etsi, 5g nr physical channels and modulation," *TS*, vol. V 15.2.0, 2018.

[106] F. Abinader, A. Marcano, K. Schober, R. Nurminen, T. Henttonen, H. Onozawa, and E. Virtej, "Impact of bandwidth part (bwp) switching on 5g nr system performance," in *2019 IEEE 2nd 5G World Forum (5GWF)*, pp. 161–166, 2019.

[107] M. Einhaus, M. B. Charaf, I. Kim, and P. Arnold, "Bandwidth part adaptation and processing time evaluation with openairinterface," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5, 2018.

[108] V. N. Ha, T. T. Nguyen, L. B. Le, and J.-F. Frigon, "Admission control and network slicing for multi-numerology 5g wireless networks," *IEEE Networking Letters*, vol. 2, no. 1, pp. 5–9, 2020.

[109] M. Kulmar, I. Muursepp, and M. M. Alam, "The impact of ran slice bandwidth subpartitioning on slice performance," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 419–424, 2022.

[110] T. Jacobsen, R. Abreu, G. Berardinelli, K. Pedersen, P. Mogensen, I. Z. Kovacs, and T. K. Madsen, "System level analysis of uplink grant-free transmission for urllc," in *2017 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, 2017.

[111] A. Lee, P. Wang, S.-C. Lin, I. F. Akyildiz, and M. Luo, "Dynamic bandwidth allocation in sdn based next generation virtual networks: A deterministic network calculus approach," in *Proceedings of the 2018 Conference on Research in Adaptive and Convergent Systems*, RACS '18, (New York, NY, USA), p. 80–87, Association for Computing Machinery, 2018.

[112] "Dynamic bandwidth part allocation in 5g ultra reliable low latency communication for unmanned aerial vehicles with high data rate traffic," *Sensors (Switzerland)*, vol. 21, pp. 1–13, Feb. 2021. Funding Information: Funding: This work was supported in part by 'The Cross-Ministry Giga KOREA Project' grant funded by the Korea government (MSIT) (No. GK18S0400, Research and Development of Open 5G Reference Model) and in part by the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) and National Research Foundation of Korea(NRF). Publisher Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland.

[113] V.-G. Nguyen, T.-X. Do, and Y. Kim, "Sdn and virtualization-based lte mobile network architectures: A comprehensive survey," *Wireless Personal Communications*, vol. 86, pp. 1401–1438, 2016.

[114] ONF, "Sd-ran white paper," *White Paper*, 2020.

[115] M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding o-ran: Architecture, interfaces, algorithms, security, and research challenges," 2022.

[116] M. Polese, L. Bonati, S. D 'Oro, S. Basagni, and T. Melodia, "Colo-ran: Developing machine learning-based xapps for open ran closed-loop control on programmable experimental platforms," *IEEE Transactions on Mobile Computing*, pp. 1–14, 2022.

[117] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5g: Ran, core network and caching solutions," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3098–3130, 2018.

[118] A. Marotta, D. Cassioli, C. Antonelli, K. Kondepu, and L. Valcarenghi, "Network solutions for comp coordinated scheduling," *IEEE Access*, vol. 7, pp. 176624–176633, 2019.

[119] 3GPP, "Study on scenarios and requirements for next generation access technologies," *TR*, vol. V38.913, 2016.

[120] 3GPP, "Nr physical layer procedures for data," *TS*, vol. V38.214, 2016.

[121] 3GPP, "Nr; multiplexing and channel coding, technical specification," *TS*, vol. V38.212, 2018.

[122] G. americas, "New services and applications with 5g and ultra-reliable low latency comunications," Nov. 2018. 5G Americas.

[123] K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, T. Balercia, S. Lee, Y. Kim, A. Ghosh, T. Thomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulos, T. S. Rappaport, G. R. MacCartney, M. K. Samimi, S. Sun, O. Koymen, S. Hur, J. Park, C. Zhang, E. Mellios, A. F. Molisch, S. S. Ghassamzadeh, and A. Ghosh, "5g 3gpp-like channel models for outdoor urban microcellular and macrocellular environments," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pp. 1–7, 2016.

[124] D. Maryopi and A. Burr, "Few-bit csi acquisition for centralized cell-free massive mimo with spatial correlation," 04 2019.

[125] A. Bana, L. Sanguinetti, E. de Carvalho, and P. Popovski, "Outage analysis of downlink URLLC in massive MIMO systems with power allocation," *CoRR*, vol. abs/2007.04784, 2020.

[126] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, aug 2019.

[127] T. K. Nguyen and T. H. Nguyen, "Performance of assigning pilot sequences in cell free massive mimo under sinr constraints," in *2018 IEEE Seventh International Conference on Communications and Electronics (ICCE)*, pp. 121–126, 2018.

[128] Y. Zhang, H. Cao, P. Zhong, C. Qi, and L. Yang, "Location-based greedy pilot assignment for cell-free massive mimo systems," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, pp. 392–396, 2018.

[129] H. Liu, J. Zhang, X. Zhang, A. Kurniawan, T. Juhana, and B. Ai, "Tabu-search-based pilot assignment for cell-free massive mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2286–2290, 2020.

[130] H. Liu, J. Zhang, S. Jin, and B. Ai, "Graph coloring based pilot assignment for cell-free massive mimo systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9180–9184, 2020.

[131] W. H. Hmida, V. Meghdadi, A. Bouallegue, and J.-P. Cances, "Graph coloring based pilot reuse among interfering users in cell-free massive mimo," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2020.

[132] Y. Li, Y. Chen, H. Huang, and X. Jing, "On massive mimo performance with a pilot assignment approach based on hungarian method," in *2016 16th International Symposium on Communications and Information Technologies (ISCIT)*, pp. 560–564, 2016.

[133] E. Björnson and L. Sanguinetti, "Scalable cell-free massive mimo systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

[134] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive mimo systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1086–1100, 2021.

[135] M. Attarifar, A. Abbasfar, and A. Lozano, "Random vs structured pilot assignment in cell-free massive mimo wireless networks," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6, 2018.

[136] S. Chen, J. Zhang, J. Zhang, E. Björnson, and B. Ai, "A survey on user-centric cell-free massive mimo systems," 2021.

[137] Y. Zhang, H. Cao, and L. Yang, "Max-min power optimization in multigroup multicast cell-free massive mimo," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2019.

[138] S. Buzzi and A. Zappone, "Downlink power control in user-centric and cell-free massive mimo wireless networks," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2017.

[139] o. T. Demir and E. Bjornson, "Joint power control and lsfd for wireless-powered cell-free massive mimo," *IEEE Transactions on Wireless Communications*, vol. 20, no. 3, pp. 1756–1769, 2021.

[140] E. Bjornson and P. Giselsson, "Two applications of deep learning in the physical layer of communication systems [lecture notes]," *IEEE Signal Processing Magazine*, vol. 37, pp. 134–140, sep 2020.

[141] S. Chakraborty, E. Björnson, and L. Sanguinetti, "Centralized and distributed power allocation for max-min fairness in cell-free massive mimo," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 576–580, 2019.

[142] T. H. Nguyen, T. K. Nguyen, H. D. Han, and V. D. Nguyen, "Optimal power control and load balancing for uplink cell-free multi-user massive mimo," *IEEE Access*, vol. 6, pp. 14462–14473, 2018.

[143] Y. Zhang, H. Cao, M. Zhou, L. Li, and L. Yang, "Power optimization in cell-free massive mimo with non-ideal hardware transceiver," *Chinese Journal of Electronics*, vol. 29, pp. 190–198, 01 2020.

[144] Y. Zhang, H. Cao, M. Zhou, S. Wu, and L. Yang, "Rate maximization for cell-free massive mimo with low-resolution adcs," in *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pp. 897–900, 2019.

[145] C. D'Andrea, A. Zappone, S. Buzzi, and M. Debbah, "Uplink power control in cell-free massive mimo via deep learning," 2019.

[146] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," 2018.

[147] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-efficient power control in cell-free and user-centric massive mimo at millimeter wave," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 651–663, 2019.

[148] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive mimo with zero-forcing precoding design," *IEEE Communications Letters*, vol. 21, no. 8, pp. 1871–1874, 2017.

[149] L.-N. Tran and H. Q. Ngo, "First-order methods for energy-efficient power control in cell-free massive mimo : Invited paper," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 848–852, 2019.

[150] T. Han and D. Zhao, "Energy efficiency of user-centric, cell-free massive mimo-ofdm with instantaneous csi," *Entropy*, vol. 24, no. 2, 2022.

[151] K. Senel, E. Björnson, and E. G. Larsson, "Human and machine type communications can coexist in uplink massive mimo systems," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6613–6617, 2018.

[152] J. Qiu, K. Xu, Z. Shen, W. Xie, D. Zhang, and X. Li, "Downlink performance analysis of cell-free massive mimo over spatially correlated rayleigh channels," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, pp. 122–127, 2019.

[153] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," 2017.

[154] Z. Zhang, D. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 1, pp. 213–225, 2020.

[155] G. E. Monahan, "State of the art—a survey of partially observable markov decision processes: Theory, models, and algorithms," 1982.

[156] J. Hu and M. P. Wellman, "Nash q-learning for general-sum stochastic games," *J. Mach. Learn. Res.*, vol. 4, p. 1039–1069, dec 2003.

[157] X. Guo, A. Hu, R. Xu, and J. Zhang, "Learning mean-field games," 2019.

[158] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[159] A. Feriani and E. Hossain, "Single and multi-agent deep reinforcement learning for ai-enabled wireless networks: A tutorial," 2020.

[160] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," *CoRR*, vol. abs/1509.06461, 2015.

[161] Z. Wang, N. de Freitas, and M. Lanctot, "Dueling network architectures for deep reinforcement learning," *CoRR*, vol. abs/1511.06581, 2015.

[162] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. M. O. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *CoRR*, vol. abs/1509.02971, 2016.

[163] C. Szepesvari, *Algorithms for Reinforcement Learning*. Morgan and Claypool Publishers, 2010.

[164] E. Ghadimi, F. D. Calabrese, G. Peters, and P. Soldati, "A reinforcement learning approach to power control and rate adaptation in cellular networks," 2016.

[165] A. A. Khan and R. Adve, "Centralized & distributed deep reinforcement learning methods for downlink sum-rate optimization," *CoRR*, vol. abs/2009.03033, 2020.

[166] M. Mohsenivatani, M. Darabi, S. Parsaeefard, M. Ardebilipour, and B. Maham, "Throughput maximization in c-ran enabled virtualized wireless networks via multi-agent deep reinforcement learning," in *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–6, 2020.

[167] L. Zhang and Y.-C. Liang, "Deep reinforcement learning for multi-agent power control in heterogeneous networks," 2020.

[168] Y. S. Nasir and D. Guo, "Deep actor-critic learning for distributed power control in wireless mobile networks," 2020.

[169] S. Saeidian, S. Tayamon, and E. Ghadimi, "Downlink power control in dense 5g radio access networks through deep reinforcement learning," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1–6, 2020.

[170] S. B. Janiar and V. Pourahmadi, "Deep-reinforcement learning for fair distributed dynamic spectrum access in wireless networks," in *2021 IEEE 18th Annual Consumer Communications Networking Conference (CCNC)*, pp. 1–4, 2021.

[171] Y. Xu, J. Yu, and R. M. Buehrer, "The application of deep reinforcement learning to distributed spectrum access in dynamic heterogeneous environments with partial observations," *IEEE Transactions on Wireless Communications*, vol. 19, no. 7, pp. 4494–4506, 2020.

[172] Y. Wang, X. Li, P. Wan, and R. Shao, "Intelligent dynamic spectrum access using deep reinforcement learning for vanets," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 15554–15563, 2021.

[173] A. Tondwalkar and A. Kwasinski, "Deep reinforcement learning for distributed uncoordinated cognitive radios resource allocation," *CoRR*, vol. abs/1911.03366, 2019.

[174] E. Ghadimi, F. Davide Calabrese, G. Peters, and P. Soldati, "A reinforcement learning approach to power control and rate adaptation in cellular networks," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–7, 2017.