ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Scuola di Ingegneria
Corso di Laurea Magistrale in Ingegneria Elettronica e
Telecomunicazioni per l'Energia LM

INTER-CELL INTERFERENCE COORDINATION ALGORITHMS FOR 5G NETWORKS

Tesi di Laurea in
Sistemi di telecomunicazioni

Candidato:
Lorenzo Vincenzi

Relatore:
Prof. Ing. Davide Dardari

Correlatore:
Ing. Giovanni Ferri

Correlatore:
Ing. Marina Lotti

SESSIONE UNICA
ANNO ACCADEMICO 2020–2021

Keywords

Inter-Cell Interference Coordination

Frequency Reuse algorithms

Ns-3

5G

Beamforming

Acronyms

3GPP Third Generation Partnership Project

5G 5th Generation

AMC Adaptive Modulation and Coding

AS Access Stratum

AWGN Additive White Gaussian Noise

BG Base Graph

BF Beamforming

BLER Block Error Rate

BSR Buffer Status Report

BWP Bandwidth part

CBLER Code Block Error Rate

CBS Code Block Size

CC Chase Combining

CN Core Network

CQI Channel-quality indicator

CRB Common Resource Block

CRC Cyclic Redundancy Check

CRS Cell specific Reference Signal

CSI Channel-State Information

CSI-RS Channel-State Information Reference Signal

DAS Distributed Antennas System

DCI Downlink Control Information

DL-SCH Downlink Shared Channel

e2e end-to-end

ECR Effective Code Rate

EESM Exponential Effective SINR Mapping

eNB evolved NodeB

EPC Evolved Packet Core

EPS Evolved Packet System

ESM Effective SINR Mapping

E-UTRAN Evolved - Universal Terrestrial Radio Access Network

FDD Frequency Division Duplexing

FR1 Frequency range 1

FR2 Frequency range 2

gNB Next Generation NodeB

GNU GNU's Not Unix

GNU GPL GNU General Public License

GTP General packet radio service Tunnelling Protocol

HARQ Hybrid Automatic Repeat Request

ICI Inter-Cell Interference

ICIC Inter-Cell Interference Coordination

InH Indoor Hotspot

IR Incremental Redundancy

IP Internet Protocol

L2S Link-to-system

L2SM Link-to-system mapping

LDPC Low-density parity-check code

LENA LTE-EPC Network simulAtor

LLS Link Level Simulation

LOS Line-of-sight

LTE Long Term Evolution

MAC Medium-Access Control

MCS Modulation and Coding Scheme

NAS Non-Access Stratum

NLOS Non-line-of-sight

NR New Radio

OFDM Orthogonal Frequency-Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

OSI Open Systems Interconnection

PDCCH Physical Downlink Control Channel

PDCP Packet Data Convergence Protocol

PDN Packet Data Network

PDSCH Physical Downlink Shared Channel

PDUs Packet Data Units

PDN Packet Data Network

PFS Proportional Fair Scheduling

PGW Packet Data Network (PDN) GateWay

PHY Physical Layer

PRACH Physical Random Access Channel

PRB Physical Resource Block

PS Packet Switched

PSD Power Spectral Density

PSS Primary Synchronization Signal

PUCCH Physical Uplink Control Channel

PUSCH Physical Uplink Shared Channel

QAM Quadrature Amplitude Modulation

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

RAN Radio-Access Network

RAR Random Access Response

RB Resource Block

RBG Resource Block Group

RBGs Resource Block Groups

RBs Resource Blocks

RE Resource Element

RLC Radio Link Control

RMa Rural Macro

RRC Radio Resource Control

RRM Radio Resource Management

RSRP Reference Signal Received Power

RSRQ Reference Signal Received Quality

RSs Reference Signals

RSSI Received Signal Strength Indicator

RV Redundancy Version

SAP Service Access Point

SCM Spatial Channel Model

SDAP Service Data Adaptation Protocol

SDUs Service Data Units

SE Spectral Efficiency

SINR Signal-to-Interference-plus-Noise-Ratio

SGW Serving GateWay

SLS System Level Simulation

SON Self Organized Network

SRS Sounding Reference Signal

SSB Synchronization Signal Block

SSS Secondary Synchronization Signal

TA Timing Advance

TB Transport Block

TBLER Transport Block Error Rate

TBs Transport Blocks

TCP Transmission Control Protocol

TDD Time Division Duplexing

TF Transport Format

TTI Transmission Time Interval

UE User Equipment

UEs User Equipments

UCI Uplink Control Information

UDP User Datagram Protocol

UL-SCH Uplink Shared Channel

UMa Urban Macro

UMi Urban Micro

UPF User Plane Function

VoIP Voice over IP

Contents

Introduction

In cellular systems that deploy 5th Generation (5G) networks there are many features that potentially may the total system to provide a good service to all the users that are covered by the cell. Each user is located within a certain cell and it is served by the base station in such a way that the signal received from the serving base station is the useful information, while signals captured from other cells represent a source for interference. Depending on the environment in which the 5G system is installed, the interference level may change due to the environmental characteristics which affect the signal propagation.
This work has been carried out at JMA Wireless company. JMA designs Distributed Antennas System (DAS) for cellular coverage of environments that require greater availability of radio resources, such as airports, hospitals, stadiums. Today's wireless systems enhance the traditional fan experience and are changing the way stadiums operate: the more these places become smart, the better will be the service offered to the users. Thus, this work focuses on the study and the implementation of Inter-Cell Interference Coordination (ICIC) schemes which have the aim of reduce the interference generated within a specific scenario. In particular, the environment that is taken into account is a generic stadium in which a 5G system is installed. The purpose is to study and verify the effectiveness of ICIC algorithms in this specific case in order to guarantee a better service to all the users that are covered by the 5G network.
Since ICIC can be applied based on multiple methods, the choice of this study is to exploit Frequency Reuse algorithms to enhance the system performance in such a way that the available bandwidth is divided into many parts so that different cells do not interfere with each other. Therefore, there are many Frequency Reuse algorithms in literature, so this work has also the purpose to gather some of them and to verify their efficiency applied to the specific scenario.
In addition to this, the study is extended to the application of 5G features such as beamforming, that is a characteristic of cellular networks which, in

this work, is investigated with the aim of applying it to the 5G network and to verify its efficiency.
The purpose is to study the behavior of the 5G network components and to modify, by applying such algorithms, the method by which resources are allocated to users with particular focus on the scheduling operations, which have been interfaced with ICIC schemes.
To do that, `ns-3` is utilized, a software that allows to implement and simulate 5G systems. Then, the implementation process consists of the construction of the scenario model, in which the base stations and the users are located. The network settings are modified and adapted to the requests and the model that we want to simulate. ICIC schemes are adopted to enhance the system performance and, finally, the simulation results are collected and observed with the aim of modify and verify the effectiveness of such ICIC schemes.

The Thesis is developed according to the following structure:
Chapter 1 is a summary of 5G features, from the basic aspects of 5G systems to the explanation of the ICIC adopted schemes. It is an introduction useful to the reader to comprehend the characteristics of 5G. Since the application of the ICIC schemes requires some modifications to a generic 5G system, it is worth to understand what are the main features that can be exploited. Then, as Frequency Reuse algorithms are not currently adopted in 5G systems, it has been necessary to implement all the features that make possible to apply these schemes based on the interfacing of the different devices involved in the operation. Furthermore, there has been the necessity to comprehend what are the 5G protocol stack layers involved in the operation and to modify the communications between them.
Chapter 2 introduces the `ns-3` software and all the specific features that it can simulate. There is a description of the different 5G layers implementation and of the models that are used to simulate the target scenario. It contains also a description of the resources allocation that is executed by the scheduler.
Chapter 3 presents the scenario that has been modeled inside `ns-3` and all the figures of merit that have been analyzed to evaluate the system performance. Beyond all the features offered by the software, there is the necessity to specify what are the settings selected for this case study. Then, this chapter focuses on those aspects that are relevant only for this work, despite of all the possibilities presented in Chapter 2.
Finally, Chapter 4 sums up all the simulation results obtained with different selected parameters in order to obtain the best result in terms of service offered to the users.

Chapter 1

Overview on 5G Medium-Access Control (MAC) and physical layers

1.1 New Radio (NR) Overview

5G is designed to support Packet Switched (PS) services to provide Internet Protocol (IP) connectivity between User Equipment (UE) and the Internet. The overall architecture consists of both the Radio-Access Network (RAN) and the Core Network (CN). The NR development within 3GPP was initiated in 2016 as a study item in 3GPP release 14. The first version of the NR specifications is available since the end of 2017 in release 15 and it is limited to non-standalone NR operations, implying that NR devices rely on Long Term Evolution (LTE) for initial access and mobility. The non-standalone NR utilizes the LTE core network, known as Evolved Packet Core (EPC). In fact, in this case, LTE and EPC handle functionalities like connection set-up and paging. Later releases will introduce standalone operations with NR connecting to the 5G core. Thus, the LTE and NR radio-access schemes and their corresponding core networks are closely related.
Moreover, even though there is the possibility for the 5G core network to handle both NR and LTE radio accesses, this work will deal with the option for which the NR radio access network is handled by the EPC.

As shown in figure 1.1, the protocol stack, subdivided into user plane and control plane, has multiple layers that interact with each other.

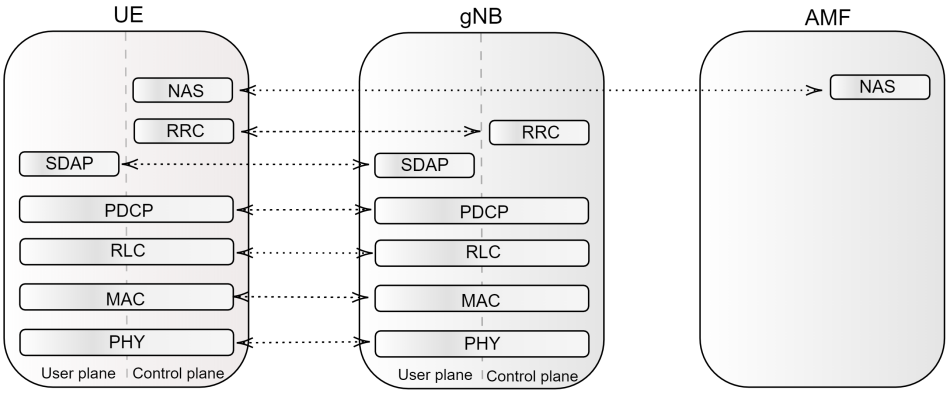

Figure 1.1: User-plane and control plane protocol stack.

The main features that will be presented in the following sections regard the MAC and the Physical Layer (PHY) layers. The first handles multiplexing of logical channels, hybrid-ARQ retransmissions, and scheduling and scheduling-related functions. The MAC communicates with Radio Link Control (RLC) through *logical channels*.
The physical layer handles coding, decoding, modulation, demodulation and typical physical-layer functions. The services offered to the MAC layer are in form of *transport channels*.

However, for completeness, the other protocol entities of the user and control plane are summarized and described below:

1. RLC is responsible for segmentation and retransmission handling;

2. Packet Data Convergence Protocol (PDCP) performs IP header compression, ciphering and integrity protection. For a device, there is a PDCP entity for each radio bearer configured[1].

In addition, the Service Data Adaptation Protocol (SDAP), which belongs to the user plane, is responsible for mapping QoS bearers to radio bearers, according to the requirements. As it will be then clear, this entity is not present in the LTE core but, for completeness, it must be remembered. Then, in the control plane, the Radio Resource Control (RRC) operates between the RRC located in the Next Generation NodeB (gNB) and the one located in the UE. It is responsible for a lot of functions, like broadcasting system information

[1] A radio bearer is an IP packet flow with a definite Quality of Service (QoS). Depending on the services requested by the UE, one or more bearers can be established.

necessary to the single device to communicate with a cell, mobility functions, and measurement reporting; the Non-Access Stratum (NAS) includes authentication, security, and the assignment of IP addresses to devices.

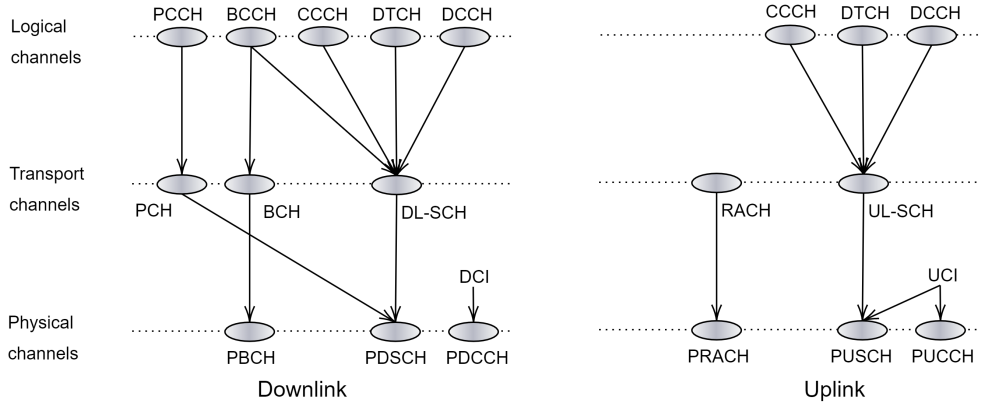

Figure 1.2: Mapping between logical, transport, physical channels.

For what concerns the MAC layer, data on transport channels are organized into Transport Blocks (TBs). Only a transport block, with its dynamic size, is transmitted over the radio interface for each Transmission Time Interval (TTI)[2]. Each Transport Block (TB) has a Transport Format (TF), which specifies how to send the TB over the radio interface: it includes information related to the transport-block size, the Modulation and Coding Scheme (MCS) and the antenna mapping. The MAC layer is responsible for the multiplexing of different logical channels and the assignment of a logical channel to a specific transport channel. The main downlink and uplink transport channels, as in figure 1.2, are Downlink Shared Channel (DL-SCH) and Uplink Shared Channel (UL-SCH), respectively.
Furthermore, there are multiple MAC control elements: scheduling-related MAC control elements, such as buffer status report, and Channel-State Information (CSI) reporting and Sounding Reference Signal (SRS) transmission are used to assist the scheduling.

As regards the physical layer, it is the responsible for mapping the signal to the appropriate physical time-frequency resources.

[2]Transport blocks may be passed down from the MAC layer to the physical layer once per Transmission Time Interval (TTI), where a TTI is 1 ms, corresponding to the subframe duration.

A physical channel consists of the time-frequency allocated resources for a transport channel, which is, in turn, mapped to its corresponding physical channel (figure 1.2). The figure shows that there are physical channels without a corresponding transport channel, and are used for control information, like Downlink Control Information (DCI) and Uplink Control Information (UCI), where the former supplies to the device the necessary information for reception of downlink data, and the latter is necessary to the scheduler for what concerns the information about the situation at the device.
In particular, the Physical Downlink Shared Channel (PDSCH) is the main physical channel utilized for data transmission, while Physical Downlink Control Channel (PDCCH) is used for downlink control information, like DCI, necessary for the reception of PDSCH. For the uplink, the Physical Uplink Shared Channel (PUSCH) is analogous to the PDSCH, as it is responsible for data transmission, while Physical Uplink Control Channel (PUCCH) is used to send Hybrid Automatic Repeat Request (HARQ) acknowledgments, which indicate whether the TB was correctly received or not.

The time-frequency resources are mapped in a way similar to LTE. Orthogonal Frequency-Division Multiplexing (OFDM) is a suitable waveform, thanks to the ease of exploiting the two domains when defining the structure for different channels and signals.
NR implements a flexible OFDM numerology with subcarrier spacings from 15 kHz to 240 kHz. The maximum available bandwidth is limited to 400 MHz and, for each frequency band, there is a subset of supported numerologies. Frequency range 1 (FR1), which goes from 0.45 to 6 GHz, only supports numerologies 0, 1 and 2, while Frequency range 2 (FR2), which has a range between 24.25 and 52.6 GHz, supports numerologies 2, 3 and 4.
Furthermore, NR provides bandwidth adaptation, that is the possibility for the UE to use a modest bandwidth when monitoring control channels, and to utilize a wide bandwidth when receiving data. NR defines *bandwidth parts* as the bandwidth over which a user is receiving transmissions of a certain numerology, so that for a UE is possible to handle that differentiation.
The selection of numerology is a balance between the overhead due to the use of subcarrier spacing and the cyclic prefix length. A larger subcarrier spacing reduces the impact of phase noise and errors. On the other hand, for a certain cyclic prefix length, the overhead increases the larger the spacing. In fact, increasing the distance between subcarriers keeping constant the cyclic prefix provides a higher overhead. Hence, a smaller cyclic prefix would be preferable. For example, for LTE the choice of 15 kHz subcarrier spacing and a cyclic prefix of 4.7 μs is a good balance, as the carrier frequency was originally designed for outdoor cellular systems that provided

Numerology	Subcarrier Spacing (kHz)	Useful Symbol Time (μs)	Cyclic Prefix (μs)
0	15	66.7	4.7
1	30	33.3	2.3
2	60	16.7	1.2
3	120	8.33	0.59
4	240	4.17	0.29

Table 1.1: Subcarrier spacings supported by NR.

up to 3 GHz carrier frequency. The NR supports a wider range of carrier frequencies, from sub-1GHz up to mm-waves. Thus, having a single numerology for all the possible scenarios is not efficient or even possible, so that for high frequencies, limitations such as phase noise are more critical and larger subcarrier spacings are needed, but at lower frequencies a large cyclic prefix is instead needed. Hence, for these type of deployments a higher spacing and a shorter cyclic prefix are suitable. NR supports a scalable numerology, with values reported in table 1.1.

For the necessity of coexistence with LTE, the NR slot structure for 15 kHz is identical to the LTE subframe structure. In the time domain, the transmissions are frames of length 10 ms, divided into 10 subframes of length 1 ms. Each subframe contains 14 OFDM symbols, whose duration is set by the numerology shown in table 1.1. Hence, with respect to the adopted numerology, the slot has a different duration (figure 1.3).

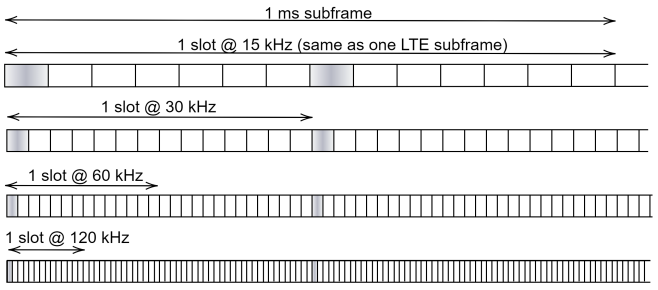

Figure 1.3: Subframes and slots in NR.

Furthermore, a Resource Element (RE) is the combination of a OFDM symbol and a subcarrier, as it is the smallest assignable physical resource. A

Resource Block (RB) is a set of 12 consecutive subcarriers. The concept of RB is different from LTE: as in NR the frequency span between subcarriers depends on the adopted numerology, a resource block with a certain numerology can occupy the same frequency of another RB with a different one. This is what flexibility of NR is all about: the gNB can allocate resources using Resource Blocks (RBs) of different formats, depending on the selected numerology.
From the perspective of the frequency domain, it is worth to note that NR is designed to support bandwidth up to 400 MHz for a single carrier, while NR devices are not able to handle such wide carrier. Then, they see only a part of the available spectrum, and that part could not be centered on the carrier, so that their DC subcarrier do not match with the center of the carrier (this is not the same in LTE, where devices see all the carrier bandwidth and coincide with its center). Hence, NR exploits also the DC subcarrier when transmitting data.

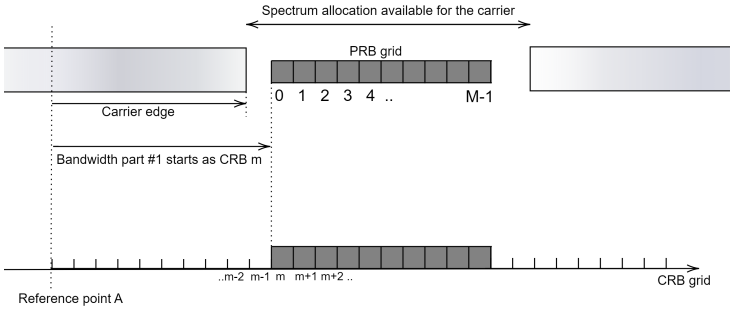

Figure 1.4: Common and physical resource blocks.

Moreover, there is a distinction between Common Resource Block (CRB) and Physical Resource Block (PRB), as figure 1.4 illustrates.
Reference Point A is a common reference point and coincides with the subcarrier 0 of CRB 0 for every subcarrier spacing. The CRBs represent the element which compose a bandwidth part, while PRBs consists of the actual transmitted signal. The figure shows that for a certain numerology, physical resource block 0 is located m resource blocks from point A, or corresponds to CRB m.

In this way, a bandwidth part is characterized by a certain numerology, and a set of consecutive blocks that start from a defined CRB.

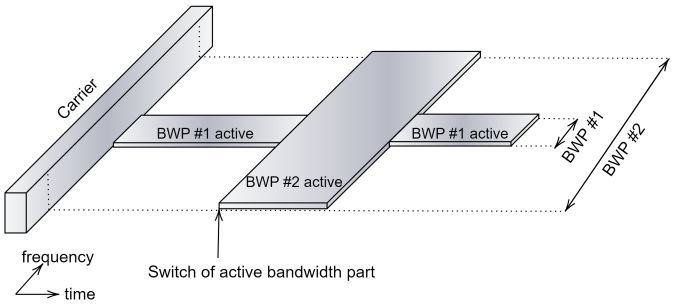

Figure 1.5: Bandwidth adaptation - Example with bandwidth parts.

A device can be configured to support up to four bandwidth parts in both downlink and uplink. This feature provides receiver-bandwidth adaptation and supports the devices, which are not able to receive the entire carrier bandwidth (figure 1.5).
On a serving cell, at a given instant there are one active downlink bandwidth part and one active uplink bandwidth part: since in downlink a device is able to receive data only inside the Bandwidth part (BWP), then the numerology of PDCCH and PDSCH are limited to the numerology configured for the specific bandwidth part. The same goes for the uplink. Thus, a device can only receive one numerology, as there is only one active bandwidth part at a time.

Regarding the radio channel, there is the need for a more detailed knowledge about the condition of the channel over which signals are transmitted. The knowledge of information about the state of the channel can be acquired in different ways: it is possible to receive measurements either from the receiver side or the transmitter side. Devices can measure the downlink channel conditions and report them to gNBs.
Alternatively, the network can acquire channel characteristics by estimating itself uplink messages. The *channel sounding* is the method by which specific signals are transmitted in order to measure and estimate channel characteristics of interest.
NR provides two different kind of signals: Channel-State Information Reference Signal (CSI-RS) and Sounding Reference Signal (SRS).
In particular, as regards CSI-RS, it occupies a single resource element and one slot in time domain; it is sent in downlink as a reference for devices, with whom can estimate channel information. Rather, SRS is an uplink sig-

nal transmitted to the gNB in order to compute itself an estimation of the channel status.
Hence, each NR device can report different measurements to the network. As regards physical layer, there is a set of possible report measurements calculated by the devices on the basis of CSI-RS or Synchronization Signal Block (SSB). The latter is the NR signal that is "always-on", transmitted with a large periodicity and used by devices which search for a cell. It consists of two parts, the Primary Synchronization Signal (PSS) and the Secondary Synchronization Signal (SSS). As it will be explained in the next chapter, the UE measurements may be based on both SSB and CSI-RS. Therefore, the device can send information like the Channel-quality indicator (CQI) and the Reference Signal Received Power (RSRP)[3].
The type of report can be periodic, semi-persistent, or aperiodic. The former is done on the PUCCH, while the semi-persistent can exploit either PUCCH or PUSCH. The latter is triggered by means of DCI signaling and takes advantage of resources on PUSCH. Chapter 2 will explain more specifically how UE measurements have been exploited in order to adapt the scheduler resource allocation to the specific channel condition of each user.

1.2 Scheduling

The scheduler is an entity that resides in the MAC layer of the gNB. It distributes radio resources to various cell users based on data present in gNB's buffers (in downlink) and Buffer Status Report (BSR) in uplink[4].
As the MAC layer is the one that performs multiplexing and demultiplexing between logical channels and transport channels, it builds MAC Packet Data Units (PDUs) (called TBs) when gNB must transmit the Service Data Units (SDUs) received from logical channels and it recovers the SDUs from the PDUs received from the transport channels[5].
The main purpose of the scheduler is to satisfy the requests of as many users as possible, taking into account the QoS requirement of each. The allocation of resources can be made following different requests, therefore, with different

[3]RSRP is a key quantity to measure and report as part of higher-layer Radio Resource Management (RRM). However, NR also supports layer-1 reporting of RSRP.

[4]The BSR is a message sent by the user to signal the necessity to transmit data in uplink.

[5]SDUs are the input data that must be processed by a specific protocol entity, PDUs are the output data.

methodologies. Scheduling algorithms are closely related to Adaptive Modulation and Coding (AMC) and HARQ. This is because the channel state measurements performed by the UEs and provided to the gNB serve to adapt the MCS. Therefore, the resources designated to each user are modulated and encoded in a different and adaptive way, based on what the user is able to receive. The MCS depends on error probability, as it defines how many useful bits can be transmitted per resource element. It defines the code rate and the modulation: the higher the CQI, the higher the modulation order and the code rate.
In addition, the concept of queue, which affects throughput and delay, depends strictly on the HARQ protocol and on the size of TBs. In fact, due to the presence of high noise or interference, there may be data received in error. Despite the MAC, RLC and PDCP layers all offer retransmission functionalities, the former targets very fast retransmissions thanks to hybrid-ARQ. HARQ performs retransmissions in two possible ways: Chase Combining (CC) and Incremental Redundancy (IR). Depending on whether the retransmitted bits must be identical to the previously transmitted or not, with incremental redundancy, which is used in NR with LDPC codes, different sets of coded bits are generated to create different Redundancy Version (RV), so that the initial transmission should include at least all the systematic bits and some parity bits, while in retransmissions parity bits which were not in the first transmission are not included. In this way, the devices that failed the first reception can rely on these different versions.
Scheduling algorithms make use of the information reported by the users, which allow gNB to manage the different requests and ensure sufficient resources to meet the QoS of each UE. Such algorithms are not standardized, that is, there is a wide range of *vendor-specific* algorithms, developed within the gNB, designed to optimize specific scenarios, together with the various operators present on the market.
In resource allocation, one of the limiting aspects is the presence of Inter-Cell Interference (ICI), especially for all those users who are on the edge of the cell. Thus, it is necessary that the scheduler includes an ICIC element, so as to take into account disturbances from adjacent cells and dispose of the resources with a different approach.

1.3 ICIC - Frequency Reuse Algorithms

In Orthogonal Frequency Division Multiple Access (OFDMA) systems, interference is a collision between RBs (figure 1.6), which has an impact on the Signal-to-Interference-plus-Noise-Ratio (SINR) associated with each col-

liding resource. The goal of ICIC is to reduce the collision probabilities and to enhance the perceived SINR by each UE.
To do this, there are several options, that include the possibility either to allocate different subsets of resources or to employ reduced power for users of different types. The aim is to utilize both these two methods, so that the users can receive a better signal quality.
Users, which are located in different positions, can be classified in two main groups: *center* User Equipments (UEs) and *edge* UEs.
The former are all those users that are located near the gNB, such that the signal quality received is higher than the quality perceived by the latter, which are located far from the gNB.
Since the study and implementation of Frequency Reuse algorithms in NR is not standardized yet, all the following mentioned studies are related to LTE.
As written in [1], the ICIC improvements are bounded to the scheduling algorithm used, the scenario in which coordination is applied, and trade-offs between the usage of resource blocks and the under-utilize of radio resources.

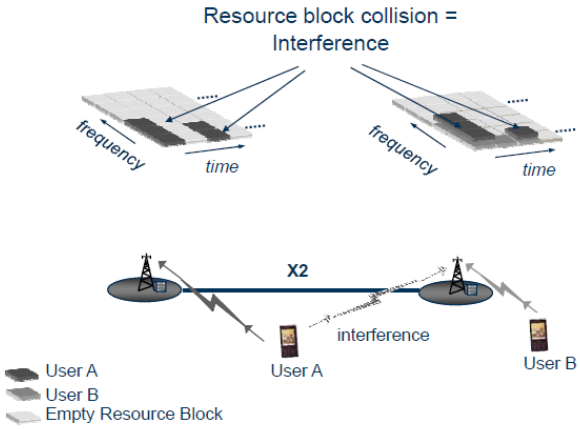

Figure 1.6: Resource Block collision - Interference ([1])

Interference can be distinguished between *inter-cell* and *intra-cell* interference. The former stands for the disturbance on frequency channels that are common to different cells, while the latter inherits the interference present on frequency channels within the same cell. As in OFDMA subcarriers are orthogonal, *intra-cell* is highly reduced.
In this work, the assumptions are that one or more RBs can be assigned to a single user at a time, and that each RB is allocated exclusively to one UE at

a time within a cell. Still, different cells can assign the same RB to different users.

In [2] ICIC techniques are distinguished into *mitigation* and *avoidance* techniques. The former include both interference randomization, where some cell-specific scrambling, interleaving, or frequency-hopping, and interference cancelation, where the interference signals are detected and subtracted from the desired received signal. The interference *mitigation* encompasses adaptive beamforming too: antennas can dinamically change the radiation pattern depending on interference level.
Beamforming is one of the principal improvements of 5G networks, therefore, apart from interference *avoidance* techniques, also beamforming is applied to the case study that will be presented in the following chapters.
Interference *avoidance* schemes can embrace both static and dynamic techniques. The first consist of frequency reuse algorithms, the second treat a Cell Coordination-based approach, that will not be adopted in this work. Static allocation schemes are not adaptive to meet dynamic demand changes, as they adapt to the cell load only by changing power used over different subcarriers. Once the allocation algorithm is set, it is difficult to modify it such that it can change the assigned resources according to the load.

As described in Chapter 2, the software used offers built-in frequency reuse algorithms shown in the figures below: for each of these, the frequency reuse is planned to restrict or allocate certain resources and power levels among users in different cells. By observing figures 1.7 and 1.8 one can demonstrate that there is a clear trade-off between available resources and interference. If all the cells, three cells in this case, are using the same spectrum resources (NoOp[6] algorithm), allocating each available RB to a user, the interference among neighbouring gNBs is the highest. On the other hand, utilizing only limited resources is not efficient. In each specific environment there are some constraints that regard the QoS that must be guaranteed to users, and the balance could hang to more resources rather than less interference.
If the two possibilities are either the use of the same resources for each cell or the enhancing of the SINR by a reuse factor higher than 1, the choice is restricted.

With regard to frequency reuse algorithms, the state of the art proposes other schemes of reuse, which are, depending on the specific scenario, more

[6]NoOp is the term used by the software to indicate that no frequency partitioning is performed between gNBs.

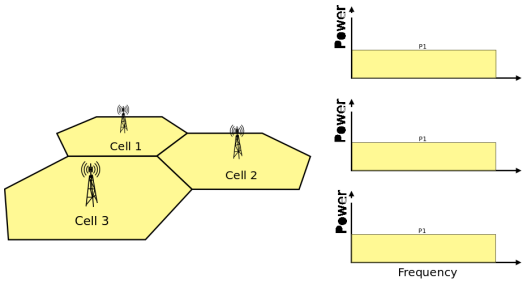

Figure 1.7: Frequency Reuse Algorithms - NoOp ([3]).

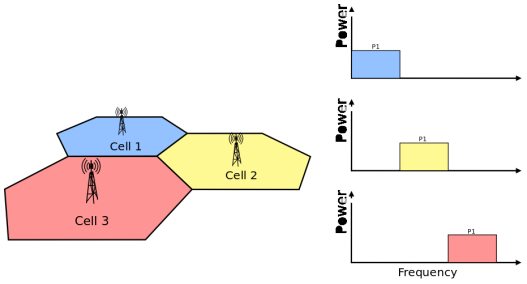

Figure 1.8: Frequency Reuse Algorithms - Hard ([3]).

oriented to either the reduction of interference or the enhancing of available resources for each cell.
In figure 1.9, which illustrates the Strict Frequency Reuse, the three cells are reserving a different edge sub-band to *edge* UEs. For all *center* UEs the resources assigned occupy the same frequency slots. In other terms, if the smallest assignable unit is a resource block, *center* UEs of different cells receive RBs at the same frequencies, while *edge* UEs do not perceive any interference, thanks to the fact that not all the available bandwidth is reserved to users. For each *edge* UE the reserved sub-band is not used by by the edge sub-bands of neighboring cells.
The power allocated to the two types of users is different: the signals received by UEs located near the gNB should be weaker, so that the interference generated is lower but the power received is sufficient because of the favorable path loss. The power assigned to edge sub-bands can be higher because no inter-cell interference is generated, even though always within the maximum allowable levels.

Figure 1.10 shows the two implementations of the Soft Frequency Reuse algorithm. The term Soft Reuse is due to the fact that the effective reuse distance can be adjusted by the division of powers between the frequencies

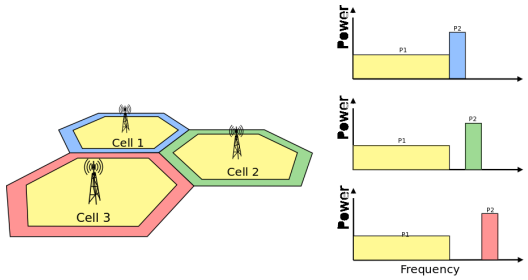

Figure 1.9: Frequency Reuse Algorithms - Strict ([3]).

used in the center and edge bands. Here, all the bandwidth is taken into account by each gNB, so the interference measured by *edge* users comes from the neighbouring cells and the current cell for the 1st version, and it comes only from neighbouring cells in the second case.

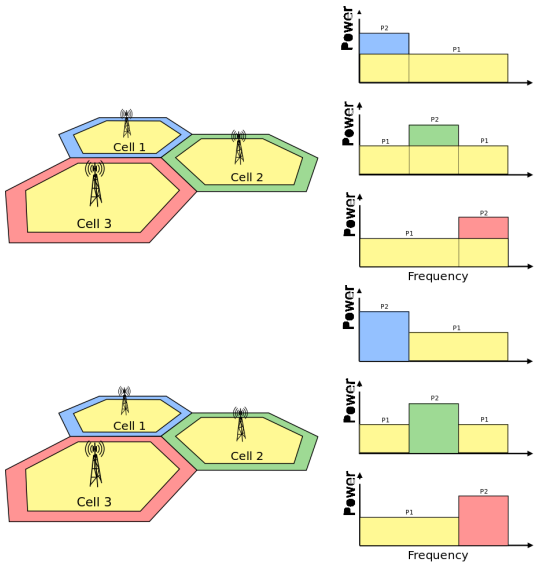

Figure 1.10: Frequency Reuse Algorithms - Soft (ver.1 and ver.2) ([3]).

In [4] the performance of static ICIC schemes are evaluated on a specific 3-cell environment, and the role of the elements that affect the spectral efficiency versus fairness is investigated. In particular, these elements are:

- Threshold to classify users, that can be *Class Proportionality* rather than *Bandwidth Proportionality*. With the former, the SINR thresholds[7] are selected so that each class (center or edge) has the same

[7]Users are classified on a channel quality basis. The following chapter will underline

number of users. With the latter, the thresholds guarantee that the number of users per class is proportional to its allocated bandwidth;

- Power level allocated to different classes of UEs;
- Number of user groups, as the increase of user groups limits the available bandwidth;
- Inter-class interference, that is caused by using the same frequencies for different user classes into different cells. Generally, center UEs receive ICI of type *inter-class* (caused by users of the neighbouring cells) and *intra-class* (caused by users of the same cell, which receive the same resources). On the other hand, edge UEs perceive only *intra-cell* ICI.

The results of the study show that *bandwidth proportionality* offers a better fairness, while *class proportionality* brings a better spectral efficiency despite of fairness, due to the fact that the number of edge users is equal to the center users, but the RBs reserved to edge users are not divided equally.
Furthermore, the higher the difference between the power assigned to the different users classes, the higher the level of fairness and the lower value of efficiency: this is caused by the fact that a small difference implies that users with worse channel conditions do not perceive a sufficient power level from the serving gNB, so that only some users are served. However, it will be proven that this does not apply to every environment, that is, each scenario has its specific requirements in terms of power assigned to different sub-bandwidths.
Then, a small number of users groups increases the efficiency, rather than a high number, which, however, ensures a better fairness.
Finally, as already mentioned, *inter-class* interference can be removed at the cost of a reduction of the entire allowable bandwidth, as illustrated in the figures above.

The following chapters will focus on these aspects and test on a specific environment the use of these Frequency Reuse Algorithms providing results consistent with the created scenario.

1.4 Beamforming

A key feature for NR is to support a large number of steerable antenna elements for transmission and reception, as NR channels and signals are designed to support beamforming.

the signals and the measurements involved.

The use of multiple antennas can offer important benefits in a mobile communication system. The use of multiple antennas has the purpose to provide diversity against fading, thanks to the fact that the channels experienced by different antennas my be partly uncorrelated.
By adjusting both the phase and the amplitude of each element, antennas at the transmitter side can provide directivity, that is, to focus the overall transmitted power in a certain direction: this is the basis concept of beamforming.
Such methods can increase the achievable data rates thanks to the enhanced power that reaches the receiver side. At the same time, multiple receive antennas can be used to provide receiver-side directivity, focusing on the reception in the direction of a target signal, while reducing the impact of interference arriving from other directions.
Furthermore, the presence of multiple antennas provides spatial multiplexing, which means transmission of multiple layers in parallel using the same resources in terms of time and frequency.
For NR the multi-antenna transmission and reception is a critical factor because of the high frequencies deployed.
Radio communication at higher frequencies involves higher propagation loss. However, this is due to the assumption that the dimensions of the receiver antenna scale with the wavelength. Then, if the carrier frequency increases, the physical dimensions of the receiver antenna are reduced, and this corresponds to a reduction in the energy captured by the antenna.
But if the antenna size is kept unchanged, with an increasing carrier frequency, the area on which the energy is captured remains the same. However, there is a correlation between the antenna size, the carrier frequency and the antenna directivity, as the latter is proportional to the physical antenna area normalized with the square of the wave length. This implies that the gain can be obtained if the receive antenna is well oriented towards the target signal.
In a mobile-communication system the use of fixed highly directional antennas is not applicable. However, the same method can be adopted by using a panel of many small antennas, which extends the overall receive antenna area.
In this case, both the dimension of each antenna element and the distance between antenna elements are proportional to the inverse of the carrier frequency. Then, for increasing frequencies, if the overall size is kept unchanged, the compensation is done by increasing the number of the small antenna elements.
Hence, the advantage of having a panel rather than a single antenna is that the direction of the transmitted or received beam can be adjusted only by

adjusting the phase of each signal applied to each antenna element.

On a high level, within the overall physical transmitter chain, the multi-antenna processing can be of two different types: it can be applied within the analog part of the chain, that is, after digital-to-analog conversion, or it can be applied in the digital part (before the digital-to-analog conversion).
The difference is that for the latter there is the need for one digital-to-analog converter per antenna element. As higher frequencies require a large number of antenna elements, the analog multi-antenna processing is the most common case. For complexity reasons, this includes per-antenna phase shifts providing beamforming.

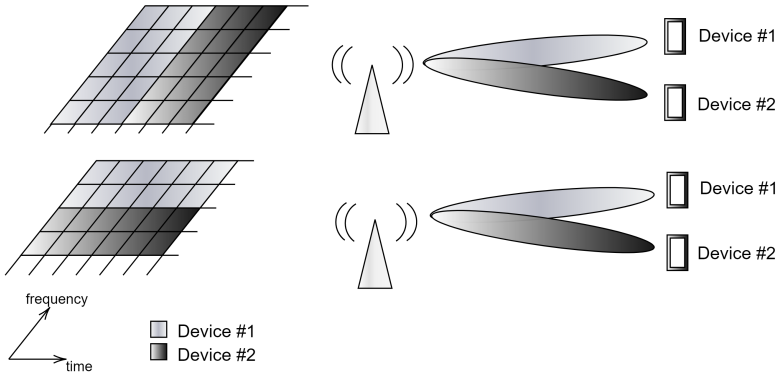

Figure 1.11: Non-simultaneous beam forming in multiple directions (up); Simultaneous beam forming in multiple directions (down).

Moreover, analog processing implies that beamforming is carried out on a per-carrier basis and it is not possible to frequency multiplex beam-formed transmissions. Hence, transmissions to different devices located in different positions must be separated in time(figure 1.11). In other words, it is not possible to allocate a different beam to each subcarrier.
On the other hand, digital processing allows for independent multi-antenna processing for multiple signals within the same carrier, that is, simultaneous beam-formed transmission to multiple devices is allowed. Thus, within one carrier it is possible to schedule multiple users with different beams.
Likewise, on the receives side, the differences between analog and digital processing are that the former implies that the receiver beam can only be directed in one direction at a time, while the latter is more flexible as it enables simultaneous beam-formed reception of multiple signals. Nevertheless, the drawback is always that the requirement is of one analog-to-digital converter

per antenna element.

The ultimate task of beam management is to establish a suitable *beam pair*, that is, a transmitter-side beam direction which corresponds to a receiver-side beam direction providing good connectivity.
The beamforming is relevant for both downlink and uplink operations. In many cases, a suitable beam pair for downlink is also a suitable beam pair for uplink and vice versa. In 3GPP this is referred to as *beam correspondence*.

In general, beam management can be divided into three parts: *beam establishment*, *beam adjustment* and *beam recovery*. The first involves all the procedures by which a beam pair is initially established; beam adjustment compensates for movements of the device and gradual changes in the environment, while beam recovery is necessary when rapid changes disrupt the current beam pair.
The following chapters will deal with all these aspects by applying all these NR features to a specific scenario, on which the results will be reported and analyzed.

Chapter 2

ns-3 Network Simulator

2.1 ns-3 Overview

In this work, 5G features and characteristics are performed and implemented by means of the ns-3 discrete-event network simulator. Ns-3 is a free software whose goal is to develop a preferred, open simulation environment for networking research: its proposal is that it should be aligned with the simulation needs of modern networking research and should encourage community contribution, peer review, and validation of the software itself.
The ns-3 project is committed to building a simulation core that satisfies the needs of the entire simulation workflow, from simulation configuration to trace collection and analysis.
The ns-3 simulation core supports research on both IP and non-IP based networks, although the large majority of its users focuses on wireless/IP simulations which involve models for LTE.

LTE-EPC Network simulAtor (LENA) is a GPLv2[1] network simulator based on ns-3 which allows LTE small and macro cell vendors to design and test Self Organized Network (SON) algorithms and solutions. Target applications for LENA include the design and performance evaluation of DL and UL Schedulers, Radio Resource Management Algorithms, Inter-cell Interference Coordination solutions, Load Balancing and Mobility Management, and other functions related to LTE. The development of LENA is open to the community in order to foster early adoption and contributions by industrial and academic partners.

[1]The GNU General Public License (GNU GPL), or simply GPL, is a series of free software licenses that guarantee end users the four freedoms to run, study, redistribute and improve the software.

As regards 5G features, 5G-LENA is designed as a pluggable module to `ns-3` (NR module). The following presented features belong to a 3GPP-compliant NR module, able to provide `ns-3` simulation capabilities in bands above and below 6 GHz, on the basis of 3GPP NR Release-15, described in [5].
Moreover, the NR module relies on higher layers and core network based on `ns-3` LTE module, thus providing the already mentioned non-standalone version of 5G networks.
This chapter will mainly focus on `ns-3` structure and methodologies which are used to implement 5G features. In particular, as the software foresees the usage of `C++` as programming language, the use of its related terms will be taken as acquired.

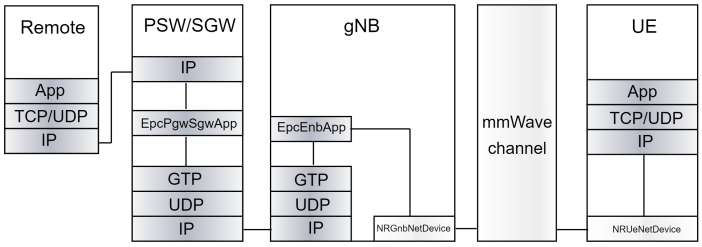

Figure 2.1: End-to-end overview.

Figure 2.1 shows how the NR module is designed to perform end-to-end simulations of 3GPP-oriented cellular networks.
The figure represents the remote host (on the left) connected to a PDN GateWay (PGW)/Serving GateWay (SGW)[2]. The packets are encapsulated through the General packet radio service Tunnelling Protocol (GTP) and sent to the gNB by the means of an IP connection, which is the backhaul[3] of the NR network. Then, when the packet is received by the gNB, it is decapsulated and transmitted through the channel by the class `NrGnbNetDevice`, which is responsible for the management of the lower layers of gNB devices. Since there is the plan for a non-standalone network, the LENA version is utilized to provide connections to higher layers and the core network, while, as already seen, the NR module is limited to realize the lower layers of the

[2]The PGW is responsible for IP address allocation for the UE and for the filtering of downlink user IP packets into the different QoS-based bearers. Through the SGW the packets are transferred to users.

[3]The backhaul portion of the network comprises the intermediate links between the core network and the small subnetworks at the edge of the network.

gNB and UE devices. In fact, figure 2.1 represents how the two classes `NrGnbNetDevice` and `NrUeNetDevice` implement the various layers.
In particular, the NR module sets up Service Access Point (SAP)[4] provider and user interfaces for the connection with the LTE upper layers. The `NrGnbMac` and `NrUeMac` classes are responsible for the SAPs that allow the MAC and PHY layers to communicate with the upper RLC layer.
It is also worth to remember that the MAC layer contains the scheduler, which implements the SAP for the LTE RRC layer configuration (`LteRrc` class) and, for the purpose of this work, for the class that designs Frequency Reuse algorithms. Furthermore, PHY classes are used to perform the downlink and uplink communications through data and control channels.
The class `NrSpectrumPhy`, presented in the following section, represents the channel and it is shared between the uplink and downlink parts.
The blocks `NRGnbBwpM` and `NRUeBwpM` manage the use of bandwidth parts. These are two virtual layers which are used to divide the bandwidth and to assign a PHY and MAC layer to each different bandwidth part. In fact, the figure shows also the possibility to have, thanks to these classes, more PHY and MAC layers configurations (each one with its own numerology, center frequency and bandwidth), as specified in the previous chapter. The NR module allows to transmit and receive flows in different bandwidth parts by either assigning each bearer to a specific one or distributing the data flow among different bandwidth parts. This feature is supported by these two managers.

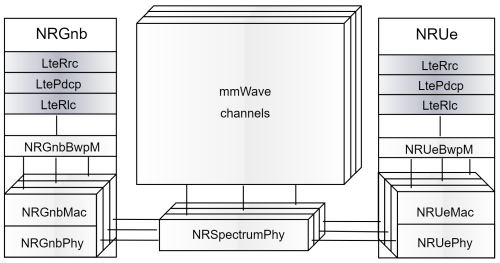

Figure 2.2: RAN overview in LENA `ns-3`

[4]A SAP is an identifying label for network endpoints used in Open Systems Interconnection (OSI) networking. The SAP is a conceptual location at which one OSI layer can request the services of another OSI layer.

2.2 PHY layer

The NR module models the numerology-dependent slot and OFDM symbol granularity to support a realistic simulation.
`Ns-3` provides an event scheduling feature by which it is possible to model time advancement. For the transmissions of frames the processing follows a logical order that involves the MAC, the scheduler and finally returns the control to the PHY. In the physical layer, allocations are extracted and, for each assignment, a new event is inserted in the simulator. From time 0 the events are inserted and processed. The events represent the advancement of time.
The frame structure is defined by the physical layer, which adopts the already defined concept of numerology (table 1.1).
As regards duplexing schemes, the NR simulator supports both Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD).
For TDD, three different slot types are defined: DL-only, UL-only and Flexible. Since one frame contains 10 slots (which, in turn, are formed by 14 OFDM symbols), for DL-only and UL-only the slots are all designed only for downlink and uplink symbols, respectively, while flexible slots have a certain number of downlink symbols, a guard band, and a certain number of uplink symbols.
In this case, the first and the last OFDM symbols are reserved to downlink and uplink control, respectively (that is, DCI and UCI signals), while the remaining symbols can be dynamically allocated between downlink and uplink symbols. In the case of DL-only slots, the first symbol is reserved to downlink control. For UL-only slots, the last slot is the one reserved to UCI.
Furthermore, since the gNB can manage multiple bandwidth parts, it is able to model FDD through the usage of two paired bandwidth parts, where one is designed for the transmission of downlink data and control, and the other one for uplink data and control. The user can configure each bandwidth part with a DL-only or UL-only pattern and then configure a linking between these two for the correct flow of control messages. In fact, it is necessary that the two parts share information such as HARQ feedback: the feedback is uploaded in uplink, but it applies to the downlink BWP, so that the configuration must ensure that messages are correctly routed from one bandwidth part to the other.
Moreover, the FDD model supports the pairing between bandwidth parts configured with the same numerology.

In both schemes the time starts at the beginning of the slot. Hence, the gNB PHY gets the allocations made by the MAC layer for the specific slot

and processes them one by one. The first symbol is reserved to downlink control, so the gNB starts transmitting it. Then, the UE can receive symbols thanks to the fact that it knows the type of the slot and the number of symbols that are reserved for the control. Through the acquisition of this information the UE can receive the DCI, from which it decides either to schedule multiple TTI to receive data or to send data. Moreover, in uplink, at the end of the slot there is a moment in which the UE can transmit its UCI. When the slot is over, another one is scheduled in a similar fashion.

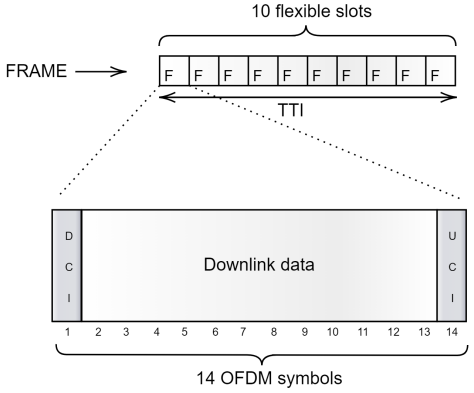

Figure 2.3: Flexible slot structure.

The software provides two types of power allocation: in the first one the power for RB is fixed, so that if a certain resource block is not transmitted it does not generate interference and the remaining power is not spread over the other RBs. The second one consists of a uniform power allocation over the active set of RBs, but it is not fixed over different slots. Then, if a RB is not available in a certain transmission, the total power is redistributed among the other resource blocks, thus providing a substantial distance from the first case, especially when the scheduler interfaces with Frequency Reuse algorithms, which make some RBs not available. This aspect will be better explained in the following chapters.

`Ns-3` includes different ways for representing the channel characteristics, following the 3GPP specifications.
In [6] it is defined how `ns-3` models the signal propagation through the wireless channel. This operation is handled by the abstract classes `SpectrumPhy` and `SpectrumChannel`.
Devices communicating through the same wireless channel have their own

`SpectrumPhy` instances, which create the Power Spectral Density (PSD) of the transmitted signals. The different `SpectrumPhy` instances are attached to the same `SpectrumChannel` object which dispatches the transmissions among the devices.
At each transmission, `SpectrumPhy` calls the `SpectrumChannel::StartTx` method which notifies the receiving device and calculates the corresponding PSDs of the transmitted signals. To account for the power attenuation and fading due to the propagation of the signal, `SpectrumChannel` relies on two standard interfaces, i.e., `PropagationLossModel` and `SpectrumPropagationLossModel`. The former models slow fading, in which the loss is constant over the frequency band of the signal, while the latter is used for fast fading models, which introduce frequency-selective losses.
These `ns-3` propagation loss models describe the signal attenuation between the transmitter and the receiver as a function of the 3D positions and the carrier frequency (path loss) and the statistical characterization of the attenuation due to the presence of obstacles between the transmitter and the receiver (shadowing). The small scale fading, that is, fast fading, accounts for the signal phase and amplitude variations due to small changes in the spatial separation between the transmitter and the receiver, and for the Doppler effect introduced by a moving terminal. While the large scale propagation effects are considered to be constant within the frequency band of interest, the small scale fading has a frequency-selective behavior, thus introducing a gain which varies within the band.
The 3GPP Spatial Channel Model (SCM)[5] can be divided into four main components:

1. Channel condition models, that are used to determine if the channel state is Line-of-sight (LOS) or Non-line-of-sight (NLOS);
2. Propagation loss models, which include path loss and shadowing;
3. The fast fading model;
4. The antenna model.

First of all, the necessary first step for the knowledge of the channel is to determine the LOS or NLOS channel condition. In [7] 3GPP defines channel models that `ns-3` develops in five different classes, each one standing for five different environments, such as Rural Macro (RMa), Urban Macro (UMa),

[5]The scope of the 3GPP SCM is to develop and specify parameters and methods associated with the spatial channel modelling that are common to the needs of the 3GPP organizations.

Urban Micro (UMi), Indoor Hotspot (InH), split into the Open Office and the Mixed versions. Furthermore, for each environment, the `ns-3` simulator can establish either the channel state is LOS or NLOS by setting it manually.
As regards path loss, [7] defines equation 2.1, where d is the 3D distance between the two endpoints, f_c is the carrier frequency, while A, B, C and X coefficients are model parameters, which take different values depending on the propagation conditions, such as the current scenario and the LOS/NLOS channel state. In particular, the path loss in dB is expressed as:

$$PL \;=\; A\; log_{10}(d) \;+\; B \;+\; C\; log_{10}(f_c) \;+\; X \qquad [dB] \tag{2.1}$$

To include the path loss and shadowing model defined in [7] the class `ThreeGpp PropagationLossModel` is implemented in order to extend the already mentioned `PropagationLossModel` and to handle the computation of the mean path loss and the shadowing component.
Since the model parameters may vary depending on the scenario, the class is extended by developing subclasses RMa, UMa, UMi, and InH, which are calculating these coefficients on the basis of 3GPP specifications.
The simulator also takes into account fast fading and antenna array models: class `ThreeGppSpectrumPropagationLossModel`, which extends class `SpectrumPropagationLossModel` interface, handles the computation of equation 2.2, where $\overline{S}_{rx}(t, f)$ is the time-varying PSD of the received signal, which depends on the transmitting and receiving beamforming vectors ($\overline{w}_{rx}$ and $\overline{w}_{tx}$), the PSD of the transmitted signal and the time-varying channel matrix $\boldsymbol{H}(t, f)$ in the frequency domain, that is,

$$\overline{S}_{rx}(t, f) = \overline{S}_{tx}(t, f)\overline{w}_{rx}^{T}\boldsymbol{H}(t, f)\overline{w}_{tx} \tag{2.2}$$

The channel matrix contains as many rows and columns as the number of transmit and receive antenna elements. It is generated by the superposition of different clusters, representing groups of multipath components that arrive and/or depart the antenna arrays with certain angles. The multipath components impact the receiving array with different delays, and the power will be scaled according to a delay-based profile.
The channel matrix generation procedure provides the possibility to select different models and parameters depending on one of the above mentioned scenarios, designed and standardized by 3GPP.
As regards interference, the powers of interfering signals are summed up together to determine the overall interference power, so that the useful and

interfering signals, as well as the noise power spectral density, are processed to calculate the SINR.

A fundamental feedback for this work is the CQI feedback, which is one of the information of which the CSI is composed. UE reports CSI to the base station using either PUSCH or PUCCH. The gNB uses the CSI to support its downlink transmissions on the PDSCH and PDCCH. The CQI is an information that can be used within link adaptation algorithms, i.e. when the base station is selecting UEs to receive resource allocations and when it is identifying the MCS.
Link adaptation, which comprises Adaptive Modulation and Coding and power control, is the term that indicates the matching between the modulation and coding scheme, and other parameters based on power control, to the condition of the radio link. AMC dynamically adjusts the transmitted information data rate to match the prevailing radio channel capacity for each user.
If a Proportional Fair Scheduler[6] is used, then UE reports high CQI values relative to their average CQI values are more likely to be selected for resource allocation. The most important concept related to CQI is that link adaptation is more likely to allocate high throughputs to UE reporting high CQI values. Thus, CQI represents a signal interference plus noise ratio rather than a signal strength.
In response to the CQI feedback, the gNB can select between QPSK, 16QAM, 64QAM and 256QAM schemes and a large range of code rates.
Three tables are defined in [8], where tables 5.2.2.1-1 and 5.2.2.1-2 are supported by the simulator. Each table is associated to one MCS table: [8] defines three tables also for MCS and each one is used on the basis of channel conditions, so that one can be preferred, with respect to Link Adaptation, depending on the channel status. In this way, it is necessary for the scheduler to receive the CQI.
CQI, which is periodically transmitted, is computed based on the the PDSCH and it is obtained from a SINR measurement. This measurement is mapped into a CQI index by the AMC module inside the scheduler.

Since the CQI index is the way for the scheduler to choose the better MCS, it also may be necessary to improve the channel quality, such that the CQI index can be higher as well as the throughput allocated to each user. This work proposes one of the possible ways to enhance channel quality, by the

[6] A Proportional Fair scheduler works by scheduling a user when its instantaneous channel quality is high relative to its own average channel condition over time.

application of Frequency Reuse algorithms. This request is supported by Measurement Reports.
The RAN controls mobility by ordering the user to perform handover to another cell. To do this, UEs periodically report their status by sending a Measurement Report message. This report contains Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ).

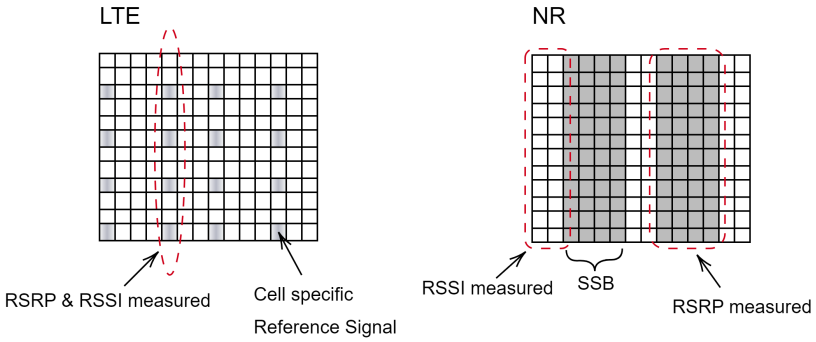

Figure 2.4: Comparison between LTE and NR measurements for RSRQ.

As already said, there are two types of signals, which would be used regardless of the measurements, that can be used for the estimation of channel quality. The first is the CSI-RS, while the second is the SSB, that encapsulates PSS and SSS, from which the user can calculate the received power. Since NR provides these two possibilities, the choice is to base Measurement Reports either on CSI-RS or SSB.
For each resource block there are resource elements dedicated to the transmission of a reference signal, that is CSI-RS, so that the receiver is able to calculate the received signal power based on a reference value[7]. Then, RSRP is the average power received from a single resource element allocated to the CSI-RS. This parameter is fundamental to calculate RSRQ, which is expressed as:

$$RSRQ = \frac{RSRP}{RSSI/N} \tag{2.3}$$

where N is the number of resource blocks across which the Received Signal Strength Indicator (RSSI) is measured, i.e. RRSI/N defines the RSSI per resource block, and it represents the total received power from all sources

[7]The structure of CSI-RS is described in detail in [8].

including interference and noise. The RSRP and RSSI are both measured across the same set of RBs.

However, figure 2.4 highlights that if SSB is used for the computation of RSRP it makes less sense to measure the RSSI from the same symbols as the RSRP because all subcarriers across the measurement bandwidth are fully occupied by the SSB. Thus, 3GPP has specified the measurement method for RSRQ such that the RSRP and RSSI can be measured across different symbols.
As shown in figure 2.4, it LTE only uses the first type of signals. Then, due to the fact that `ns-3` does not implement handover for NR, i.e. there is not a feature that provides a periodic Measurement Report by UEs, it is simpler, from a complexity point of view, to adopt CSI-RS, that are allocated in a way similar to LTE Cell specific Reference Signal (CRS).
Periodically (every 200ms), the physical layer of each UE sends a Measurement Report to the gNB containing the RSRQ calculated. The MAC scheduler, which makes use of Frequency Reuse algorithms, uses this information to classify the users so that, once a threshold has been set, it is possible to define if the user is a Center User or an Edge User.
However, since Measurement Report is related to the handover procedure, it must be triggered on the basis of some criteria defined in [9]. In particular, the RRC layer, which is the mainly involved entity in this operation, implements Section 5.5.4.1 “Measurement report triggering - General” of [9]. It takes into use the latest measurement results and evaluate them against an entering condition: an applicable entering condition will lead to a new reporting, so measurement reports would be produced and submitted to gNodeB. The entering condition is satisfied when, among the handover Events, Event A1 is triggered. This is triggered when the serving cell has a RSRQ higher than a certain threshold.
Since there is not the intention to use handover algorithms, this step is unnecessary to the aim of this work. Hence, the software has been modified in order to always trigger a Measurement Report, which is fundamental to classify users. Chapter 3 will show how the threshold can be designed in order to obtain different results applied into a specific scenario.

The simulator also includes a PHY abstraction model for error modeling which is consistent with NR specifications, including Low-density parity-check code (LDPC) coding, different MCS tables and transport block segmentation. It also supports HARQ based on IR or CC.
The error model of the data plane is developed according to Link-to-system mapping (L2SM) techniques. The performance evaluation of a wireless com-

munication system that is based on computer simulation requires a division of simulation levels. The simulation levels are Link Level Simulation (LLS) and System Level Simulation (SLS). They require an interface between them, that is Link-to-system (L2S) interface. It transfers the link level performance characteristics and directs it to the system level for the performance evaluation.
L2S interface is designed to capture the link level performance based on the SINR versus Block Error Rate (BLER), since these measures are commonly used for the assessment of the link level performance for the PHY layer of wireless systems. The L2S interface has to consider the physical layer technology such as modulation mapping, channel coding and radio access interface. Since the transport block is the payload which is passed between the MAC and PHY layers undergoing PHY processing at the transmitter before being mapped onto the PDSCH for transmission over the air interface, in OFDM multi-carrier transmission, for a certain block there are many different SINR values due to the multipath fading experienced by that transport block, which is allocated by physical layer by means of multiple RBs. The L2SM process receives inputs consisting of a vector of SINRs of allocated RBs, the MCS selection, the transport block size delivered to PHY and the HARQ history. The error model has the aim of making an assessment based on multi SINRs values versus the transport block error rate. Therefore, an effective value is needed to be taken for the multi-state values of a transport block and should be compressed to only a single scalar value. This concept is called Effective SINR Mapping (ESM). The calculation of the effective value will be necessary for the computation of the block error ratio for that transmitted block over that instantaneous channel and propagation characteristic.
To do this, `ns-3` simulator uses Exponential Effective SINR Mapping (EESM) as L2S mapping function. Figure 2.5 shows how the physical layer abstraction model is implemented: it consists of SINR compression, LDPC Base Graph (BG) selection[8], code block segmentation, mapping of the effective SINR to BLER for each code block, mapping of code BLERs to the Transport Block Error Rate (TBLER).

Since the process is complex and the purpose of this work is not to modify the error model developed in `ns-3`, while the intention is to study and to take advantage of the designed features offered by the software, all the blocks which compose the chain will be mentioned keeping in mind that for a more in-depth knowledge a more detailed analysis of the elements that make up the process is necessary.

[8]BG selection is based on specifications in [10].

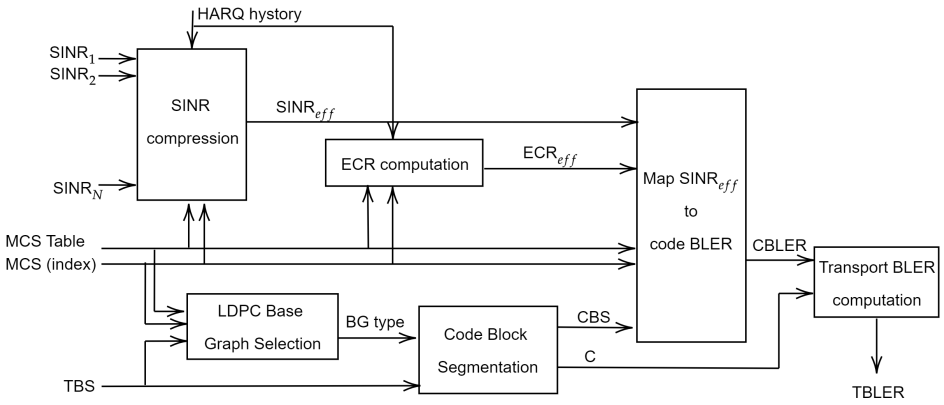

Figure 2.5: NR PHY abstraction model.

As already said, [8] defines three tables of MCSs, where Table 5.1.3.1-1 contains MCS indexes from 0 to 28, while Table 5.1.3.1-2 from 0 to 27 (Table 5.1.3.1-3 is not supported by `ns-3`). The MCS index selection is communicated to the UE by the gNB through the DCI. Each MCS index defines an Effective Code Rate (ECR), a modulation order the resulting Spectral Efficiency (SE).
The HARQ history is based on the HARQ method: if CC is used, the history contains the SINR per allocated RB, while for IR the history consists of the last computed effective SINR and the number of coded bits of each of the previous retransmissions. Then, for the former, the ECR after n retransmissions remains the same after the first transmission. SINR values are summed across the retransmissions and the combined value is used for EESM. In HARQ-IR, every retransmission contains different coded bits than the previous one. Therefore, the effective SINR and the ECR must be calculated after each retransmission.
Moreover, when the number of total bits in a TB is larger than the maximum Code Block Size (CBS) (which is 8448 bits for BG1 and 3840 bits for BG2) code block segmentation is necessary for LDPC coding. Hence, each transport block is split into a certain number of code blocks following specifications of Section 5.2.2 in [10].
The SINR compression block calculates the effective SINR deriving from SINR measures of every received RB following equation:

$$SINR_{eff} = -\beta \, ln \left(\frac{1}{v} \sum_{n \in v} exp \left(-\frac{SINR_n}{\beta} \right) \right) \tag{2.4}$$

It represents the effective SINR for the single transmission, where $SINR_n$ is the SINR value belonging to the n-th resource block, v is the set of allocated RBs, and β is a parameter that has to be optimized. In particular, given an experimental BLER measured in a fading channel with a specific MCS, the β value is designed in a way that the effective SINR of the channel approximates to the SINR that would produce the same BLER, with the same MCS, in Additive White Gaussian Noise (AWGN) channel conditions[9].
Once the effective SINR is computed, there is the need for SINR-BLER lookup tables to find the corresponding Code Block Error Rate (CBLER). These are provided by the software by performing extensive simulations using its NR-compliant link level simulator.

$$TBLER = 1 - \prod_{i\,=\,1}^{C} (1 - CBLER_i) \simeq 1 - (1 - CBLER)^C \tag{2.5}$$

Finally, if code block segmentation is used, there is the necessity to convert the CBLER found from the link-level simulator's lookup table to the TBLER for the input TBS. Following expression 2.5, the code BLERs of the C code blocks are combined to get the BLER of a transport block.

Furthermore, with regard to beamforming, PHY layer supports two types of algorithms: ideal and realistic. Since beamforming algorithms are based on Beamforming (BF) vectors, which are clusters containing the beams that a transmitter or receiver antenna uses, there are two possible ways for the transmitter to calculate them. The first is to assume the ideal knowledge of the channel, without consuming any resource to obtain information about channel conditions. The second is to explore the channel with some measurements based on SRSs. The Sounding Reference Signal is transmitted in uplink and is used by the gNB to evaluate the channel status.
This work will focus on the realistic method because of the purpose to demonstrate the effectiveness of Frequency Reuse algorithms applied to a case study in which the channel condition for users is the mainly investigated issue. Thus, it would be approximate to calculate beams according to an ideal knowledge of the channel.
For the realistic beamforming method there is a helper class named `RealisticBeamformingHelper` which implements `BeamformingHelperBase` interface. The latter creates a BF task when a UE is attached to a gNB. The BF task is composed of a pair of connected devices for which the realistic BF

[9]The calibration of parameter β is conceived according to [11].

helper will manage the update of the BF vectors. For each BF task it is created an instance of realistic BF algorithm, that is realized with class `RealisticBeamformingAlgorithm`, which is connected to `NrSpectrumPhy` SRS SINR trace. Each realistic BF algorithm instance is also connected to its helper through callback to notify it when BF vectors of device pair need to be updated. When BF vectors need to be updated, the function `GetBeamformingVector` of realistic BF algorithm is called, which calculates for each pair of predefined beams of the receiver and transmitter a channel estimate and based on that selects the best BF pair. Thus, the three parts of beam management mentioned in Chapter 1 are all handled by the beamforming helper: the establishment is done when the UE is attached to the gNB, while the adjustment and the recovery are obtained by the function `GetBeamformingVectors`. When BF vectors need to be updated, function `GetBeamformingVectors` is called and, in turn, it calls `GetEstimatedLongTermComponent` for each pair of beams of the receiver and transmitter couple in order to estimate the channel quality of each of them based on SRS reports.
Lastly, function `CalculateTheEstimatedLongTermMetric` selects the best BF pair.
The algorithm offers a predefined set of beams that may be changed by the user depending on the scenario. The computation of the best beam is made according to the channel estimation deriving from SRSs, which allow the gNB to calculate the SINR and to build a channel matrix containing representative elements of the channel status.
The BF vector trigger update is SRS count event, i.e. after every N SRS are received, the BF vectors are updated. The count event can be set by the user, involving that lower values could produce a processing time due to beamforming that would affect the performance and resulting in an unnecessary use of beamforming for many scenarios. This aspect will be treated in the following chapters, where its effectiveness will be evaluated and applied to a specific case study.

2.3 MAC layer

The retransmission mechanism is handled by the HARQ entity residing in the MAC layer. It performs fast retransmissions of each erroneously received transport block. Therefore, if there is not the possibility to decode the packet, the received signal still contains information, which would be lost by discarding incorrect received packets. HARQ with *soft combining* makes

it possible to store the packet in a buffer memory and to later combine it with the retransmission in order to obtain a single combined packet with a higher reliability. The already mentioned Chase Combining and Incremental Redundancy are the two different methods by which the HARQ performs *soft combining*.
The `ns-3` HARQ entity residing in the scheduler is in charge of controlling HARQ processes for generating new packets and managing the retransmissions both for downlink and uplink. It assumes that retransmissions of the same HARQ process use the same MCS and the same number of resource blocks. This restriction is due to the specification of the rate matcher in the 3GPP standard as in [10], in which the algorithm fixes the modulation order for generating the different blocks of the redundancy versions.
According to the HARQ feedback and the RLC buffer status, the scheduler generates a set of DCIs which include new transmissions and retransmissions of erroneously received HARQ blocks. In fact, the mechanism consists of multiple stop-and-wait protocols, each operating on a single TB. Thus, the transmitter stops after each transmission and it waits for an acknowledgment, which comes from the receiver (HARQ feedback). In addition, multiple stop-and-wait processes can work in parallel in order to enhance the throughput, that is, while the transmitter is waiting for the acknowledgment from one HARQ process it can transmit data to another HARQ process.

The NR module of `ns-3` supports variable TTI downlink OFDMA, which means that the number of symbols allocated to one user is variable, based on the scheduler, and it is not fixed as in LTE. For decoding transmissions, UE relies on a bitmask, which is an output of the scheduler and is encapsulated in the DCI[10]. The bitmask contains the list of Resource Block Groups (RBGs) assigned to that user. In NR, as specified in [8], a Resource Block Group (RBG) can group from 2 to 16 RBs, depending on the subcarrier spacing and the operational band. Depending on the adopted Frequency Reuse algorithm, transmissions will enable only a certain number of RBGs to each UE.
Then, as the main output of the scheduler is a list of DCIs for a specific slot, the enclosed information regard the transmission starting symbol, the duration (in number of symbols) and the RBG bitmask. The scheduler assigns the resources for active downlink and uplink flows and notifies the MAC layer of the decisions.
The different phases that the OFDMA scheduler follows are:

[10] As a reminder, DCIs are control signals sent on PDCCH and captured by UEs in order to obtain the essential information about PDSCH.

- BSR and CQI processing. The former brings information about the status of the user buffer, the latter is the measure relative to the quality of the channel. The MCS can be computed by the AMC for each UE based on the CQI. Information such as MCS and BSR are stored in order to be later read to determine UE capabilities;

- Calculation of the total number of active flows for uplink and downlink. This is necessary because of the two different delays which characterize uplink and downlink transmissions. The delay consists of the number of time slots that must pass between the moment in which the decision is taken and the moment in which the signal is transmitted: the default values are 2 slots for downlink and 4 slots for uplink, meaning that, for the latter, there is a prioritization with respect to the former;

- Uplink decisions are not considered for the slot indicated by the MAC layer, but for a slot in the future. Instead, the procedure for downlink allocations is started relative to the slot indicated by the MAC. Since this work will mainly deal with downlink transmissions, all the requirements and impacts of uplink operations are not treated.
  Therefore, for the priority attributed to the uplink scheduling, the number of symbols previously given for uplink data has to be considered during the downlink scheduling. Then, the available symbols can be distributed by giving priority to HARQ retransmissions and then to new data based on two scheduling levels.

The two scheduling levels are the time-domain level and the frequency-domain level. The former performs the scheduling of the symbols per beam: the scheduler selects a sequence of consecutive OFDM symbols in a slot to assign to a specific beam. The latter consists of the scheduling of RBGs per UE in the single beam, i.e. the scheduler determines the allocation of RBGs for the OFDM symbols of the corresponding beam (figure 2.6).
The first level can be performed in a load-based or a round robin fashion: the latter is a way to fairly dispose resources to all users, while the load-based regards a more dynamic method. In fact, the calculation of the load is based on the BSRs, which determine the current status of users buffers. The second level can apply different algorithms to allocate RBGs among UEs associated to the same beam. These include round robin, proportional fair and maximum rate manner.
With round robin the available RBGs are divided evenly among the UEs associated to the beam. Proportional fair scheduling distributes RBGs according to a metric that considers the actual rate, based on CQI, and the average rate that has been provided in the previous slots to the different

UEs. Maximum rate, instead, distributes RBGs among UEs according to a maximum rate metric that considers the actual rate of the various UEs.

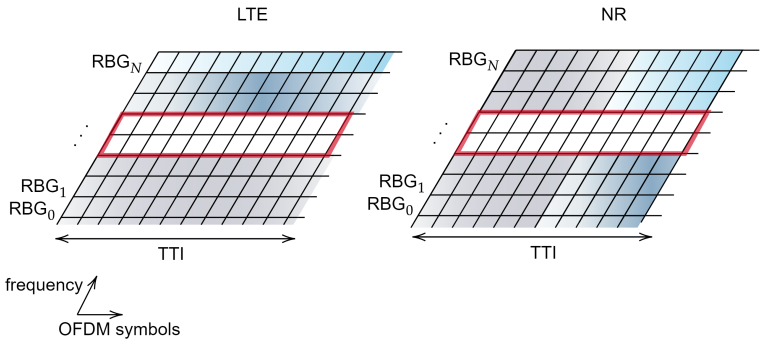

Figure 2.6: LTE (left) and NR (right) resources assignment - Resource Elements assigned to different UEs are represented with different colours. The red line highlights unusable RBGs.

However, the second level scheduling must take account of Frequency Reuse algorithms as for a specific cell not all the RBGs could be available for all users. In fact, figure 2.6 shows that the use of these algorithms may make unusable some specific RBGs.
To do this, it has been necessary to interface the scheduler with Frequency Reuse algorithms by means of specific SAPs.
Since `ns-3` did not provide Frequency Reuse algorithms to the 5G module, it has been necessary to create SAPs in order to allow the scheduler to obtain the available RBGs for each user within each TTI.
Hence, when the downlink scheduling is assigning a certain RBG, it must verify that, for the adopted Frequency Reuse algorithm, it is available for the UE.
The figure highlights the difference between LTE and NR scheduling: the former allocates all the available OFDM symbols within a TTI only to a user and, for each UE, it selects a certain number of RBGs, depending on the scheduling algorithm and their availability, as some RBGs could not be allocated to any UE inside a specific cell.
The NR improvement consists of the allocation of OFDM symbols to more than a user inside the duration of a TTI. As each user can be served within a certain beam, the scheduler calculates the number of OFDM symbols to allocate to a specific beam and, in the next level, it decides what RBGs are allowed and how to assign them to different UEs. The difference is that NR

offers a higher granularity in terms of OFDM symbols since each one can be assigned to a user independently from the selected RBG, thanks to the first level scheduling that is a time-domain scheduling.

Chapter 3

Network Settings and Performance Evaluation

3.1 Scenario and Settings

The purpose of demonstrating the effectiveness of Frequency Reuse algorithms must be accomplished through a realistic environment, so that it can be possible to extract results based on a specific case study. In this chapter we describe the scenario used to assess the performance of the algorithms proposed.
The starting point is a urban area modeled as a Urban Micro area for `ns-3`, characterized by the presence of a high interference level due to the proximity of the various gNBs. Interference derives from the fact that the same frequency resources are used within the same time instants by different gNBs.
In particular, studies are made on a generic stadium in which six sectors are considered. Each one is covered by a single gNB, which serves to the users located therein. In figure 3.1 a model is drawn, which represents the described scenario with a generic number of users within each cell, that is, one of the drawn six sectors.
As it is depicted, some UEs are located near the gNB, while some of them are far from it. The high level of interference caused by neighbouring cells makes difficult for the latter to decode messages correctly. Then, there is the necessity to enhance the quality of the signals that are received in downlink and of those which are sent in uplink.
The parameter that has to be improved is the already mentioned SINR, which allows users to perceive a better signal quality in order to reduce the packet loss.

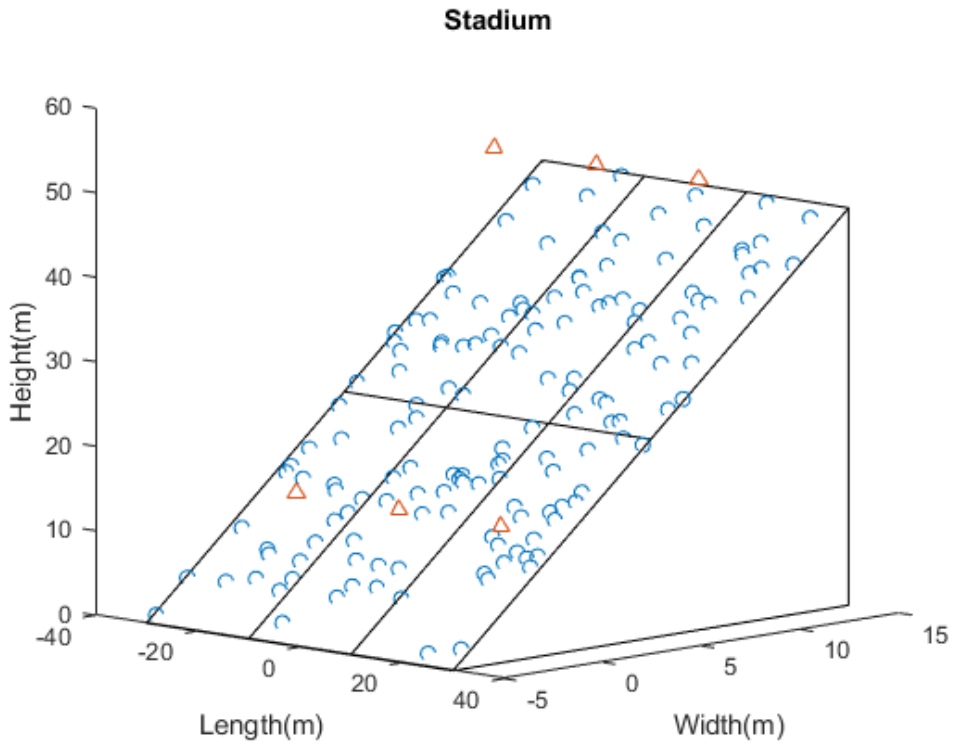

Figure 3.1: MATLAB model of the scenario. UEs are the blue circles, gNBs are the red triangles.

The main challenge is to guarantee a better QoS to all the users, focusing on those UEs which are located far from the gNB, thus increasing the signal quality level for distant users. The other important aspect to consider is the high density of users that characterizes this type of environment. One of the goals will be the possibility to use beamforming and then to exploit directive beams toward the target UEs.

The settings that are used regard all the aspects treated in the previous chapters. A basic concept is that this work starts from a generic existent system installed in the stadium which offered a LTE coverage. Then, since there are many aspects related to operators, which manage the availability of resources in terms of allowable bandwidth and sustainable data rates, it must be considered that not all the possible features offered by a 5G system can be utilized. Thus, 5G features are adapted to LTE limitations and the performance is evaluated based on the different methods and improvements offered by NR. In fact, for this specific case there are some limitations to take

into account, such as the maximum bandwidth, that for LTE is equal to 20 MHz. Hence, this work focuses on a study of the possible methods to adopt in order to optimize the system performance, regardless of the available total bandwidth, rather than to exploit all the allowable resources.
Then, the first assumption is that a TDD scheme is adopted, so that downlink and uplink operations benefit from the channel reciprocity. Thus, a single bandwidth part is established for both the communications. The central frequency is set to 2.1 GHz, inside the FR1 band, while the transmission power is 43 dBm per cell. This involves that the total power is spread over all the RBGs assigned to each UE at any TTI. Hence, for the already mentioned power allocation type, within each TTI, each allocated RBGs may either have a power level that can change depending on the number of available RBGs for that cell or keeping a fixed value in such a way that, even if a RBG is not assigned to any user, the power level of the other RBGs remains the same. This is an important aspect which must be taken into account for the calculation of the SINR related to Frequency Reuse algorithms. Therefore, by referring to figure 1.10 it can be noted that Edge users have a different power offset than Center users. In fact, [9] defines the PDSCH configurations that can be set to design different transmission methods in downlink. Thus, it is possible to assign different power offsets to Center UEs and Edge UEs. This is done with the aim of reducing the interference between neighbouring cells. By referring to Soft Frequency Reuse algorithm (version 1) depicted in figure 1.10, if the power offset, which is applied to the assigned RBGs, is set to -6 dB for Center UEs and 3 dB for Edge UEs (which are the minimum and the maximum values provided by 3GPP specifications) the result is that the bandwidth allocated to Center users, which is the largest, causes a lower interference to both Edge and Center UEs from the other cells, while the bandwidth reserved to Edge users may cause a higher interference, especially when, within a TTI, only Edge UEs are served in one cell, so that all the available power is allocated to those reserved RBGs.
Hence, the selection of the power allocation must be done very carefully.

The resource allocation type determines the number of RBs grouped into a RBG. The selected allocation type 0 designs 8 RBs per RBG. If numerology 0 is set, the total number of RBG available per TTI is calculated based on the subcarrier spacing and the width of the bandwidth part. Then, for 5G, for a 20 MHz channel bandwidth the number of PRB is 106. For 8 RBs per RBG and a subcarrier spacing of 15 kHz, the total number of RBGs is equal to 13.
The bitmap on which the RBGs assignment is based is of size 13 bits with one bitmap bit per RBG such that each RBG is addressable. The RBGs

shall be indexed in the order of increasing frequencies and starting at the lowest frequency of the carrier bandwidth part. The RBG is allocated to the UE if the corresponding bit value in the bitmap is 1, whereas the RBG is not allocated to the UE otherwise. Hence, the granularity of Frequency Reuse algorithms is based on 8 RBs assignable per user and a total number of assignable RBGs equal to 13.
Figure 3.2 shows an example of resources assignment when a Hard Frequency Reuse algorithm is applied. The RBGs available for cell 1 are the first two, coloured in blue.

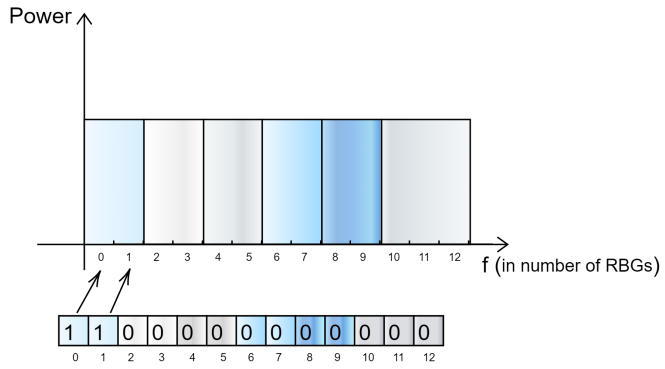

Figure 3.2: Hard Frequency Reuse algorithm - Resources assignment.

The simulation settings also regard the transmission direction, which is decided to be the downlink. Even if this work focuses on downlink data transmissions, all the results can be extended and compared also with uplink communications.
For each user a bearer is established, so that a continuous flows of data is guaranteed. Then, since the 5G scheduler may handle a data rate of 150 Mbps, User Datagram Protocol (UDP) traffic is generated and the total 150 Mbps per cell are spread among all the users. UDP is a transport level protocol which is combined with the IP level in order to exchange data between different nodes of a network. As it does not guarantee the retransmission of lost packets, all the functions related to retransmissions are due only to the lower layers. Moreover, the choice of UDP instead of Transmission Control Protocol (TCP) is due to the fact that `ns-3` does not provide the TCP protocol.
Hence, since UDP is more suitable for real-time applications, the goal for the scheduler is to handle operations that require a minimum bit rate, thus focus-

ing the work on communications that do not tolerate excessive packet transmission delay and can tolerate some data loss, such as Voice over IP (VoIP).

3.2 Performance evaluation

The above depicted scenario presents a high interference level which causes, among the other things, a high packet loss level. Users that perceive a lower signal quality are not able to decode received packets and need for more retransmissions. Since the TBLER level is high, which means that the erroneous decoded bits per transport block are even more, since the Incremental Redundancy version of the HARQ entity is used, the number of redundancy versions required is even higher.
This type of loss involves a reduction of the throughput. Throughput is the main parameter that will be discussed and used to evaluate the system performance. In fact, it is the ratio between packets correctly received and the total transmitted packets related to the data rate offered to the user, so that if a certain bit rate is assigned to a user, the performance is evaluated by referring to the percentage of bits per second correctly decoded based on the ratio between received and transmitted packets. Then, at the UDP level, since there are no upper layers modeled and involved in the following simulations, for each data flow the throughput is computed based on UDP packets. A data flow is provided to each user by means of a bearer. For the aim of this work, which is the evaluation of system performance based on real-time applications, the designated bearer offers a VoIP service.
The first aspect to take into account in order to enhance the throughput is the TBLER, which must be reduced. TBLER can be reduced by applying Frequency Reuse algorithms, which provide a better SINR to each user.
If the perceived SINR is higher, the probability to erroneously decode a packet is lower. Then, as shown in figure 3.3, the possibility to exploit a higher SINR involves the possibility to use a high MCS index.
As a high spectral efficiency enhances the throughput, it is preferable to use a better modulation and coding scheme, in order to increase the amount of bits per second that are sent on the channel.
The figure shows that, for a fixed value of TBLER, the higher the MCS, the higher the SINR required.
The choice of a higher MCS is related to figure 3.4. In fact, in a specific scenario a certain data rate may be required, as operators want to guarantee a certain service to users; thus, a high data rate is, in general, required. As depicted in the figure, the higher the data rate required from the EPC, the higher the MCS needed to satisfy this request.

The results represented in figures 3.3 and 3.4 are obtained with multiple simulations in which only a user is present in a single cell. The TBLER evaluation is based on a end-to-end (e2e) transmission from the gNB and the UE, for which are considered five different MCS indexes. For each MCS index the TBLER is calculated by the average TBLER obtained during 1 s of packet transmission with a 2 Mbps data rate. The TBLER stands for the average packet loss and is different depending on the MCS index. In fact, by referring to table 3.2, if MCS 0 is used, the modulation order is low, thus enhancing the probability for the user to correctly decode a packet also for low values of SINR. For the same value of SINR, MCS 28 returns the worst TBLER, as for the user is impossible to decode neither the first transmitted packet nor its redundancy versions. For instance, for a TBLER equal to 0.5 it is necessary a SINR of about 20 dB if the MCS index used is MCS 28, but it is sufficient a SINR of -8 dB if MCS 0.

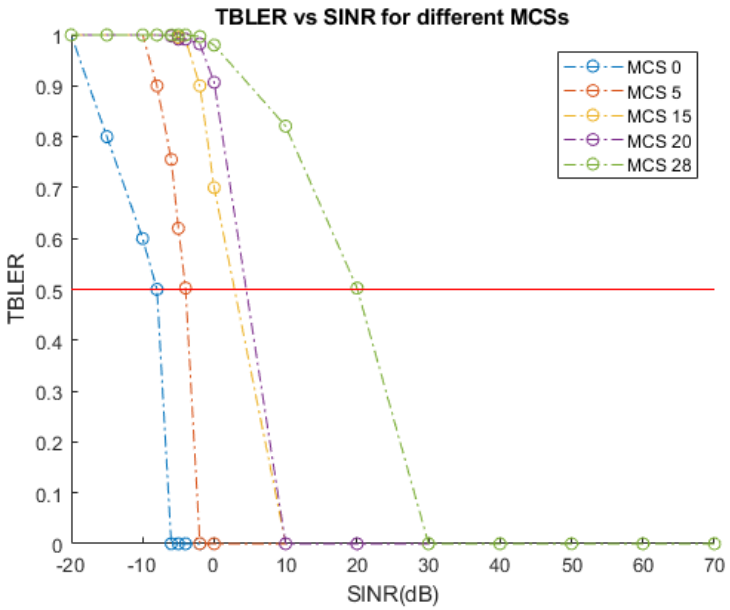

Figure 3.3: TBLER evaluation depending on SINR, for different MCS values.

Hence, the reason for the request of a higher MCS is derived from figure 3.4. The results are obtained from a simulation in which a single UE is located in a single cell, that is, a single UE covered by a single gNB, with a SINR value equal to 40 dB. The e2e communication is tested with different data rates provided by the gNB. The curves represent the throughput (only in terms of received packets on the total transmitted) obtained by multiple

MCS index	Modulation Order	Target Code Rate
0	2(QPSK)	0.117
5	2(QPSK)	0.370
15	4(16QAM)	0.602
20	6(64QAM)	0.554
28	6(64QAM)	0.926

Table 3.1: MCS Index Table 1 (64QAM Table).

end-to-end (e2e) transmissions with different data rates and variable MCS index. If a 200 kbps data rate is requested, by assuming that the TBLER is equal to zero (the SINR is ideally the highest), the throughput is equal to 100%.

If a 20 Mbps data rate is required by the user, a low modulation order and a low code rate are not enough to optimize the throughput: the simulations show that with MCS 0 the throughput is about 8.3%, but with MCS 28 it reaches the 99.6 %. Hence, the purpose is to enhance throughput by increasing the SINR perceived by each user in order to reduce TBLER, but also to exploit a high spectral efficiency with the aim of sustain a higher data rate. It is worth to note that values reported in figure 3.4 are referred to the only ratio of received packets on transmitted packets, which is not the actual figure of merit taken as a performance index. The throughput which will be taken into account includes the multiplication with the actual data rate, so that if a 100 Mbps bit rate is offered to a user with MCS 28, the 60 % of packets are correctly decoded, thus offering 60 Mbit per second. This is not equal to the case in which a bit rate of 60 Mbps is guaranteed with MCS 20, because the ratio between received and transmitted packets is 60 % but it corresponds to approximately 40 Mbit per second.

However, the design of such a scenario is not simple: the limited bandwidth involves a limited availability of resources. The transmitted bits are encapsulated into a large number of transport blocks. Then, if the requested data rate is high, there is the need for larger transport blocks for allocating more resources per TTI. In fact, as one TB is assigned per TTI to a user, the larger the TB, the higher the information encapsulated. However, a limited bandwidth supports only a limited data rate value. For example, the figure is obtained by a simulation in which the UDP packet size is equal to 500 bytes, which are sent to the lower layers to be processed and sent over the air. The maximum number of transmitted packets depends on the selected data rate:

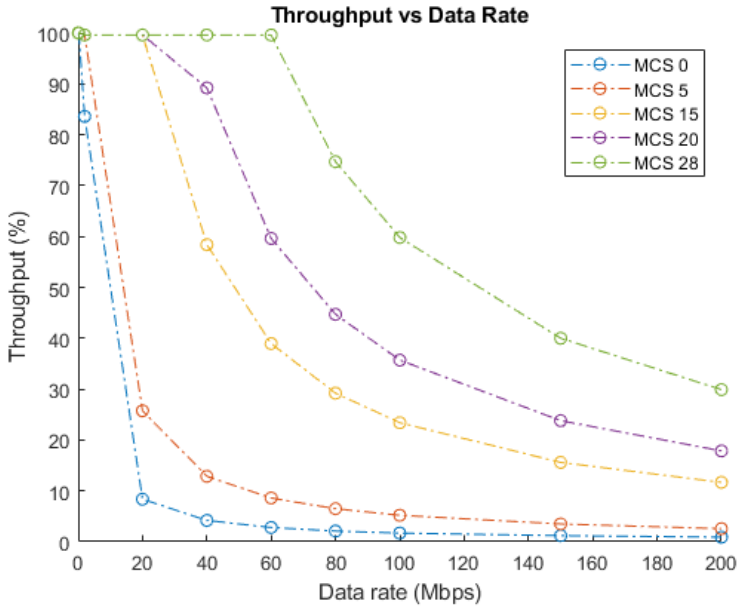

Figure 3.4: Throughput evaluation depending on data rate, for different MCS values.

e.g., for a 200 Mbps data rate, the total number of packets transmitted in 1 s is equal to 50000, but if the TBS is small, due to the limited total bandwidth, not all the packets are received, even if any packet is lost due to TBLER. This means that the scheduler needs a larger bandwidth.
The problems introduced by Frequency Reuse algorithms consist of a more limited available bandwidth for each user. Furthermore, the performance evaluation is not only focused on the application and testing of Frequency Reuse algorithm, but also on the optimization of resources allocation, which is an operation of the scheduler.
It is clear that in such a scenario it is not possible to sustain a high data rate because of the high density of users and the limited amount of resources.

In particular, Transport Blocks consist of million bits, which are processed by the physical layer before being mapped onto the PDSCH for the subsequent transmission over the air interface. As the TB is the payload which is passed between the MAC and the PHY layers through transport channels, the TBS is determined by the PHY. The channel coding processes the TB and, when designing the method by which the TB is built up, it may include code block segmentation, that is necessary when building up TBs because of the LDPC encoding executed by the PHY layer, which can handle a certain code block

size[1]. In fact, the transport channel coding executed by the PHY layer may include within its steps the code block segmentation, which consists of the process depicted in figure 3.5.

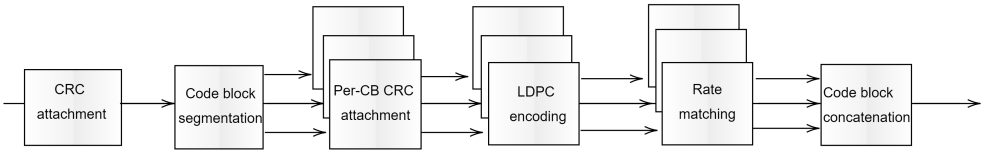

Figure 3.5: Code block segmentation.

Since 5G supports large bandwidths together with a wide range of transmission duration and variations in the overhead (that is, all the features which may be configured, such as CSI-RS), the transport block size has to be determined depending on the resources that the scheduler can allocate to a user on the PDSCH. Thus, there is not a specific table to take as a reference to assign different transport block sizes depending on the configured settings. Therefore, in Section 5.1.3.2 of [8], 3GPP reports the adopted formula-based approach combined with a table that is dedicated to the smallest TB sizes.
Furthermore, to schedule a certain number of bits the scheduler should have sufficient resources to satisfy the data rate in terms of OFDM symbols and RBGs. As already said, the scheduler assigns resources based on two scheduling levels, that is, time-domain and frequency-domain. Hence, the number of allocated Resource Elements must be sufficient to sustain such a transmission of data.
The number of information bits brought by the TB is obtained by equation:

$$N_{info} = N_{RE} \cdot R \cdot Q_m \tag{3.1}$$

where N_{RE} is the number of Resource Elements allocated within a TTI, R is the adopted code rate and Q_m is the modulation order. Based on tables 5.1.3.1-1 and 5.2.2.1-2 in [8], which provide the code rate and the modulation order associated to each MCS index, the computation of the number of

[1]The LDPC coder in NR is defined up to a certain code block size, that is 8424 bits for base graph 1 and 3840 bits for base graph 2.

information bits reflects the limitation of resources: even if the highest MCS index is used (it provides a code rate of 0.93 and a modulation order equal to 6) the number of information bits that will be encapsulated into the TB is limited by the number of the assigned REs.
The number of REs is calculated by the multiplication of the number of RBGs allocated to the user within the specific TTI (N_{RBG}), the number of RB per RBG (N_{RB}^{RBG}), the number of OFDM symbols within a RB which are reserved to data (N_{sym}^{RB}) and the useful subcarriers (N_{sc}^{RB}), that is, those which are reserved to data, following the formula:

$$N_{RE} = N_{RBG} \cdot N_{RB}^{RBG} \left(N_{sc}^{RB} \cdot N_{sym}^{RB} \right) \tag{3.2}$$

Therefore, based on N_{info} the TB size is computed following the 3GPP specifications, as depicted in figure 3.5: if the variable N_{info} is smaller than 3824 bytes there is not segmentation and the TBS is determined by consulting table 5.1.3.2-1 of [8]. A CRC is attached to the TB, which encapsulates only a code block. On the contrary, if the information bits are more than 3824 bytes there is the need for the code block segmentation in the channel coding, thus attaching a CRC to each code block, as shown in figure 3.6.

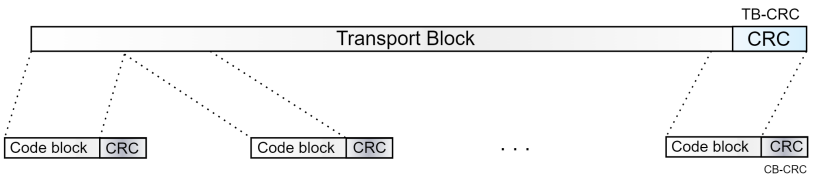

Figure 3.6: Transport block CRC.

In the case of code block segmentation the transport block CRC is also attached and it may be redundant as the set of code block CRCs should indirectly provide information about the correctness of the complete TB. However, as the code block group retransmission is a mechanism to detect errors per code block, thus providing retransmissions only of erroneous code block groups instead of the complete transport block and enhancing the spectral efficiency, the TB CRC adds an extra level of protection in terms of error detection.
Thus, the redundancy necessary for this operations adds another constraint

to take into account when considering the sustainable data rate by a gNB. Not all the bits inside the TB bring information, thus decreasing the allowable bit rate in the e2e communication.
For simplicity, by neglecting all the redundancy bits that are necessary in these operations, if it is taken as an assumption a TTI duration of 1 ms and if it is considered that, in general, within each TTI only a TB is assigned to each user, a bit rate of 150 Mbps would require 150000 information bits allocated per TTI. Then, a TB has to include all these information bits, but the variable N_{info} is limited by the number of allocated Resource Block Groups. Hence, as shown in figure 3.7, N_{info} must be at least equal to 150000, but it is limited by N_{RE}, which can be increased by expanding the total bandwidth, in order to dispose of more resources. For instance, the maximum value for the product $(N_{sc}^{RB} \cdot N_{sym}^{RB})$ is 156, that is the product of 14 OFDM symbols and 12 subcarriers. Even if all the available RBs were assigned to the user, with a total bandwidth of 20 MHz (106 RBs), and the MCS 28 index were used, the value N_{info} would be expressed by:

$$N_{info} = 0.93 \cdot 6 \cdot 156 \cdot 106 = 92270 \tag{3.3}$$

The TB could contain at most 92270 information bits. The actual number of bits included within a TB is limited, among the other things, by the fact that some bits are reserved to CRCs and the number of Resource Element reserved to data does not reach its maximum value because of the many factors that limit the resources reserved to data within a RB. Thus, the only way to have the possibility to sustain such a data rate is to expand the 20 MHz bandwidth, in order to allocate more resources.

Moreover, because of these constraints, there is the necessity to clarify the limitations of such a work. Firstly, it is worth to note that all the results that are provided in the following chapter are obtained by simulations on a software, that can not return the real performance of such a network because of the ideality of many factors. Secondly, the purpose of this work is not to focus on the effective QoS provided by the gNB to each user, but more to find a way to increase the performance in terms of TBLER, SINR and Spectral Efficiency, starting from a starting point, in which no ICIC is provided, and applying all the presented algorithms, in order to optimize the system performance with respect to the case without Inter-cell Interference Coordination.

Chapter 4 will show how Frequency Reuse algorithms can be applied to

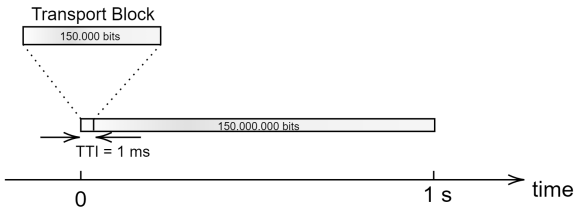

Figure 3.7: Transport block size.

the presented scenario. It will be also presented another algorithm based on the combining of beamforming and Frequency Reuse algorithms. This new scheduling algorithm depends on the two-level scheduling that can be exploited in order to assign to each beam a certain number of OFDM symbols and a certain number of RBGs to each user.

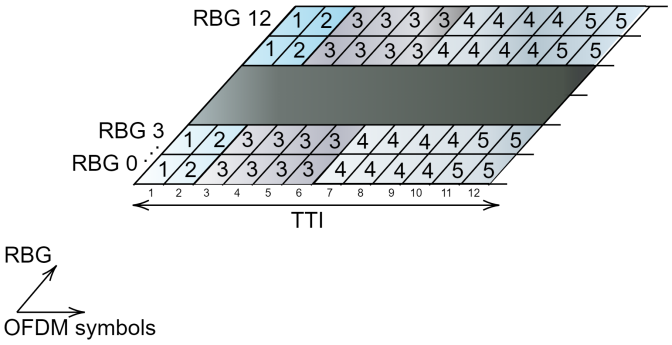

Figure 3.8: Beamforming with Frequency Reuse algorithms scheduling.

Thus, within a TTI the scheduler disposes of 12 OFDM symbols reserved to data (cfr. Section 2.2) which are spread among all the beams used by the gNB. Each beam has a number of assigned OFDM symbols. Then, a user within the cell receives these OFDM symbols only if there are sufficient RBs available. In fact, the frequency-domain level scheduling decides if, for a vari-

able number of symbols, there are enough frequency resources to assign to a specific user. For example, figure 3.8 represents a generic scheduling process, where resources are allocated to 5 different users: the three colours indicate the three different beams selected for the users. User 1 and 2 are scheduled within the same beam, as well as user 4 and 5, while user 3 does not divide the beam resources with other UEs. The scheduler selects a certain number of OFDM symbols for each beam, based on the load (which, for simplicity, is neglected in this example). The example shows that 2 OFDM symbols are selected for the first beam, 4 are dedicated to the second one, and the remaining 6 are allocated to the third beam. Then as user 1 can utilize only 1 OFDM symbol per RBG, if all the RBGs were available for that cell, the total RBGs assigned to user 1 would have been equal to 13. However, when Frequency Reuse algorithms are applied, maybe only a restricted number of RBGs are available within a cell. Hence, the total number of RE reserved to each user is limited.

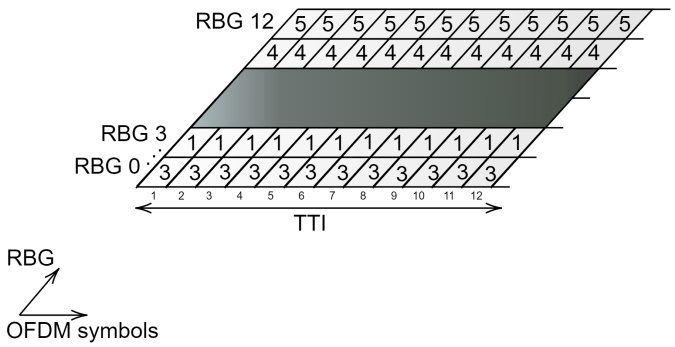

Figure 3.9: Frequency Reuse algorithms scheduling.

Moreover, this algorithm provides a significant variation to the use of Frequency Reuse algorithms. In particular, figure 3.9 shows the case for which no beamforming is used. Here, the number of OFDM symbols assigned to a user is always equal to 12 because of the absence of the first level scheduling, that divides the symbols among the different beams. Hence, the figure shows that, for the same scenario above presented, the assignment of resources is only restricted to the selection of the available RBGs.
Furthermore, the purpose of the new proposed algorithm is to limit the use of beamforming. In fact, since Frequency Reuse algorithms classify users into two main categories, which are Center UEs and Edge UEs, beamforming will

be applied only to those users which are located near the gNB. As the purpose is to enhance the SINR perceived by each UE, if reuse algorithms are applied, Edge users already have a high SINR level while, in general, Center users have a lower SINR that can be enhanced by the use of beamforming.
Therefore, since Frequency Reuse algorithms increase the SINR but limit the allowable bandwidth per user, the choice that was made is to reserve beamforming only to Center users and not to Edge users, which already have a high SINR level.
However, the beamforming operation is set when the UE is attached to the gNB, while the classification of Center and Edge UEs is an event that occurs periodically, every 200 ms, thus when the communication is already established.
For the purpose of this work, this requires an hypothesis, that is the assumption that the gNB already knows the position of the UEs. Taking this assumption as granted, then the gNB can decide either to adopt beamforming or not depending on the position of each user. Subsequently, when the connection is established, users will be classified into Center or Edge UEs. Thus, it will imply that there will be users for which beamforming will be adopted and they will be also classified as Center users by Frequency Reuse algorithms, and there will be users classified as Edge UEs for which a reserved bandwidth will be assigned and beamforming will not be used.
The decisions of these two classifications are based on two thresholds: the first depends on the distance of the UE from the gNB, so that a high value will guarantee a higher number of users for which beamforming is used due to the fact that beamforming operation is adopted only for those users whose distance is lower that the threshold. The second is the already mentioned RSRQ threshold, which is utilized by Frequency Reuse algorithms to classify users into Center or Edge UEs depending on the signal quality perceived by each one.
This implies that an optimum value must be found out for the two thresholds and, by decreasing the distance threshold, the risk is that a user which is considered as an Edge user by the RSRQ threshold could be located at a distance that is lower than the distance threshold, thus enabling beamforming also for it. As it will be explained in the next chapter, this will not cause any problem.

Figure 3.10 depicts two concentric circles which represent the two thresholds. Colour grey represents the area in which beamforming is applied, while colour blue is the area where users do not exploit beamforming. These two areas are divided by the distance threshold. The other threshold is the RSRQ one and divides users between Center and Edge UEs.

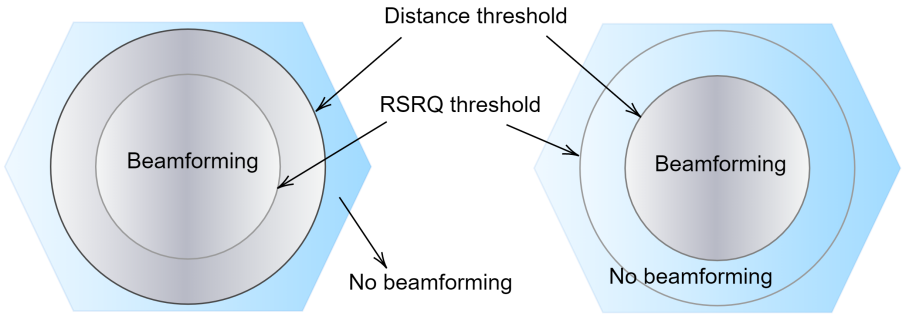

Figure 3.10: Distance threshold and RSRQ threshold.

There are two possible cases for the thresholds design: the first case involves that beamforming is used also for Edge users, while the second one causes that beamforming is not adopted for some Center UEs. Both these cases will be investigated in the following chapter, as there will be a certain optimum value for the two thresholds. It must be specified that the higher the RSRQ threshold, the higher the number of Edge UEs: if the RSRQ measured by a user is below the RSRQ threshold, it is classified as Edge UE. In terms of spatial distance, the higher the RSRQ threshold, the smaller the circle: with a high threshold all the users which perceive a high RSRQ are classified as Center UEs, while all the others are Edge users. Thus, if the RSRQ threshold is expanded from a spatial point of view, it means that the requested RSRQ is lower.

The hypothesis is necessary because the `ns-3` NR module does not provide the network attachment process. In fact, 3GPP offers the possibility to choose either to use beamforming or not when the user device is attached to the gNB, but `ns-3` configures manually the device attachment, thus bypassing all the network attachment process, which may be useful for this work.
However, the 3GPP features that can be adopted to obtain the users position by the gNB may be related to the network attachment process. The two main steps of initial access are the Cell Search and the Random Access. The first is based on SS blocks, from which users measure the RSRP and decide what gNB to connect to.
Then, once a device has found a cell it may access the cell by the random access procedure. This operation includes four steps, which involve the device transmission of a preamble, the network transmission of a Random Access

Response (RAR) that indicates the reception of the preamble and provides a time-alignment adjusting the transmission timing of the user based on the timing of the received preamble, and two final steps which have the aim of resolving potential collisions due to simultaneous transmissions of preambles by multiple users.
The random-access preamble is referred to as Physical Random Access Channel (PRACH), that is, a special physical channel, represented in figure 1.2. The transmission timing of NR uplink transmissions is controlled by the network by means of a closed-loop timing control. It consists of regularly provided time adjustments commands needed by the gNB to ensure a time alignment. This mechanism, described by 3GPP in [12], is named Timing Advance (TA) and consists of a negative offset at the device between the start of a downlink slot received by the user and the start of the uplink slot. The network can control the timing of the signals received from the users at the gNB by designing appropriately the offset for each device. For instance, devices that are located far from the gNB encounter a larger propagation delay than nearer devices, thus needing for a different offset. The timing advance value is determined by the network based on the device uplink transmissions. Before preamble transmission there is not a closed-loop timing control. The time adjustment is based on the preamble transmission timing and the received timing of the SS block. Figure 3.11 depicts the general concept of Timing Advance: a user far from the base station requires a large timing advance to compensate the larger propagation delay.

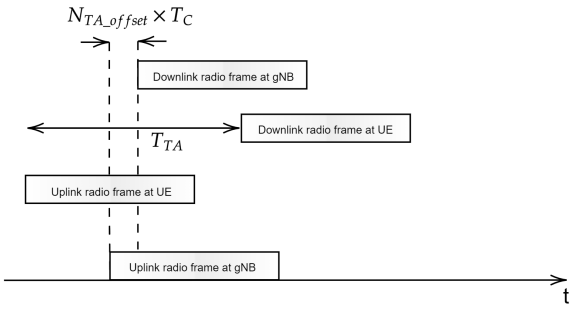

Figure 3.11: Timing Advance offsets.

The represented variables report the offsets that are set in order to design the synchronisation by means of timing advance (T_{TA}), which can be modeled

Subcarrier Spacing	μ	$N_{TA} \cdot T_C$ - Time Domain	$N_{TA} \cdot T_C$ - Equivalent Distance
15 kHz	0	520.83 ns	78.13 m
30 kHz	1	260.42 ns	39.06 m
60 kHz	2	130.21 ns	19.53 m
120 kHz	3	65.10 ns	9.77 m

Table 3.2: Timing Advance step size.

as:

$$T_{TA} = (N_{TA} + N_{TA\_offset}) \cdot T_c \qquad \text{where, } T_c = \frac{1}{480000 \cdot 4096} \tag{3.4}$$

The factor $(N_{TA\_offset} \cdot T_c)$ ensures that an uplink radio frame finishes before the start of the subsequent downlink frame. The factor $(N_{TA} \cdot T_c)$ is the timing advance step size. Therefore, it is defined a set of 12 bits which provides a value within the range from 0 to 3846. This variable corresponds to T_A and it is used to calculate N_{TA} following the formula:

$$N_{TA} = T_A \cdot 16 \cdot \frac{64}{2^{\mu}} \tag{3.5}$$

where μ is the subcarrier spacing index (0, 1, 2 or 3 for 15, 30, 60 or 120 kHz). This means that $(N_{TA} \cdot T_c)$ represents the round trip propagation delay between the UE and the gNB. Thus, the maximum value of N_{TA} defines the maximum supported cell range and the product with T_C corresponds to a step size for timing advance which is variable depending on the adopted subcarrier spacing. In fact, table 3.2 reports the different values of timing advance step size, thus including that with a fixed subcarrier spacing there is a quantization that defines the possible measurable distance. For example, with μ=0 the step size is about 520 ns, which corresponds to 78.13 m; this means that users which are located at a distance lower than 78.13 m are all seen by the gNB at the same distance. This affects the precision when the intention is to exploit this feature to measure the distance of the UE from the gNB. In the presented scenario, the average distance between the cell and the user is about 20 m. However, as reported in table 3.2, by increasing subcarrier spacing it is possible to improve the granularity of the timing advance, thus obtaining the desired results.

As regards beamforming, for 5G there is the possibility to apply beam-sweeping for SS-block transmission, that is, to transmit SS blocks in different beams with a certain periodicity. This involves that a device located in a certain downlink beam may be unaware of the SS blocks transmitted from the cell to the other beams. Beam-sweeping for the SS block transmission enables receiver-side beam-sweeping for the reception of uplink random access transmissions.
Therefore, there is the possibility to establish a suitable beam pair already during the initial access phase and to apply receiver-side beam sweeping during the network attachment for the preamble reception. This involves that, once a device has transmitted a preamble it waits for a response, which could follow the beam used for the SS block, that was acquired during the cell search. By transmitting the response using the same beam as the SS block, the device can use the same receiver beam utilized during cell search. However, for the purpose of this work, it may be possible for the gNB, once it has received the preamble[2], to decide either to utilize the same beam or not to apply beamforming for that user. This may be configurable both in a real gNB and in the `ns-3` simulator, even if this version of the NR module does not offer the related features.

The next chapter addresses the improvement brought by the proposed frequency reuse algorithm based on beamforming with respect to the baseline scenario where neither Frequency Reuse algorithms nor beamforming are applied.

[2]The preamble is not sent over a specific beam because the gNB and the UE are not connected yet.

Chapter 4

Simulations and Results

4.1 Frequency Reuse algorithms

Once defined all the settings necessary to the simulation of the e2e communication, the next step is to verify the effectiveness of Frequency Reuse algorithms applied to the scenario presented in the previous chapter. The goal is to enhance the user QoS in terms of TBLER, SINR and Spectral Efficiency. The most important figure of merit which is taken into account is the already mentioned Throughput. It is intended as the bit rate offered by the gNB to a specific user, multiplied by the ratio between received packets and the total transmitted packets.

The starting point is a scenario characterized by a high interference level due to the position of gNBs and the UEs within the cells, and due to the power level utilized by gNBs to transmit data. The following results are referred to the scenario depicted in Chapter 3, where 5 UEs per cell (that is, per sector) are randomly located within the six stadium sectors[1]. As the purpose is to guarantee a high data rate to each user, a first evaluation of the throughput is done with different bit rate values offered by the single cell. Figure 4.1 shows the throughput obtained with different data rate values when no ICIC is used, that is, NoOp algorithm presented in Chapter 1 is applied. Those values are referred to the single cell, thus if a cell covers 5 UEs, the total bit rate must be divided by all of them. In other words, if a cell has to offer a 200 kbps data rate to 5 UEs, then each one can receive, in ideal conditions and disposing of an infinite set of resources, 40 kbit per second. The single

[1]This study focuses on 5 UEs per cell because, for simplicity, it has been taken into account a small number of UEs. However, it could be demonstrated that the final results also apply with a larger number of UEs.

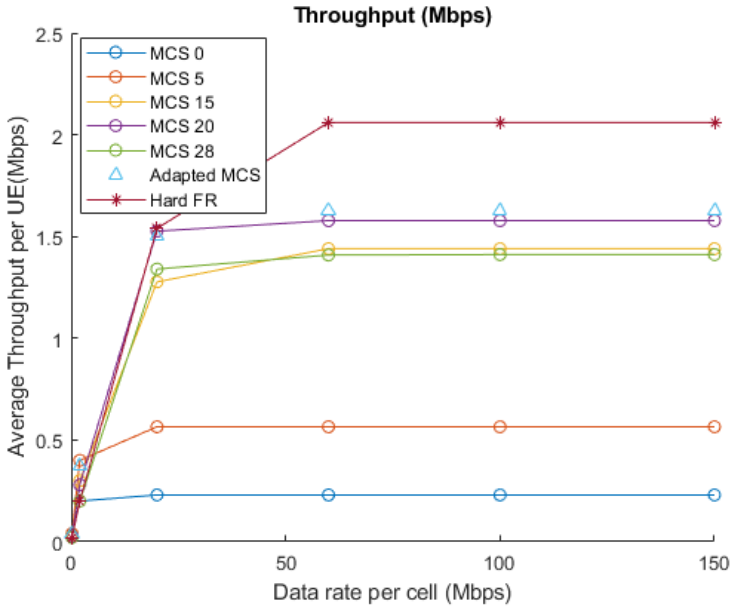

Figure 4.1: Average Throughput per UE (Mbps) with different data rates per cell - NoOp and Hard Frequency Reuse algorithms.

user throughput is intended as the received bits per second on the total bits per second offered by the gNB. Figure 3.4 shows the throughput that a single user can reach when a single cell offers different data rate values. The result is that a 60 Mbps data rate is the only one that allows to reach the maximum throughput in ideal conditions, that is, MCS 28 and zero TBLER. This is not to be intended as the maximum bit rate which can be offered by the single gNB. In fact, if a 150 Mbps data rate is exploited by the gNB, the reachable throughput is limited to the 40% but, thanks to the fact that not always the scheduler guarantees a certain level of fairness, a user can reach a throughput value which is higher than the one obtained when the gNB offers a 60 Mbps bit rate. In other words, if 5 UEs are located within each cell, one can reach 30 Mbps at the expense of another one. However, when the gNB data rate is 60 Mbps, the maximum throughput that a single user can reach is equal to 12 Mbps. This aspect does not influence the following results because all the factors that affect the system performance bring the single users to receive only a small part of the total bit rate that the gNB offers them. Thus, if a 60 Mbps data rate is selected, the results are similar to the case of 150 Mbps.
As already said, the maximum reachable throughput per user depends, first of all, on the total available bandwidth. Thus, figure 4.1 depicts the average throughput of the users depending on different data rates. The throughput is

calculated as the average throughput of all the users, where the throughput of the single user is computed by multiplying the bit rate offered by the gNB to the user and the ratio of received packets on the transmitted packets. The curves are based on a fixed MCS index: since a high spectral efficiency is required, an increasing MCS index is applied in a non-adaptive way in order to demonstrate that the curves depicted in figure 3.4 can be obtained only if the conditions are ideal. In fact, a bit rate of 60 Mbps can be reached by a user when the SINR is the highest, the MCS index is equal to 28, and when it makes use of the total 20 MHz bandwidth. In particular, figure 4.1 demonstrates that a 12 Mbps bit rate can be received by a user (the cell offers 12 Mbps to each one of the five users) only with ideal conditions, such as a high SINR and the availability of the total 20 MHz bandwidth. Since there is not an Inter-cell Interference Coordination method applied, then the SINR is low and the packet loss is high. The total bandwidth available for each user is shared among the 5 users within the cell, so that for each TTI the TBS allocated to the single user will be limited. Then, while figure 3.4 demonstrates that even with ideal conditions the maximum data rate assignable to a single user is limited by the bandwidth, figure 4.1 shows that the single user throughput, intended as received bits per second on the total bit per second offered by the gNB, is limited by the TBLER due to a low SINR, and by the limited available total bandwidth.
Therefore, by forcing the scheduler to allocate resources without the intervention of the AMC entity, thus working on the five depicted curves (MCS 0, MCS 5, MCS 15, MCS 20, MCS 28), it does not provide the expected results ensured by figure 3.4 due to resources limitation and the high level of interference. Hence, if the expected average throughput for the users were equal to 12 Mbps, as the scheduler can assign in these conditions a maximum 60 Mbps data rate per cell, this value would not be reachable due to non-ideal conditions.

The limitation of resources can be demonstrated by the application of the first Frequency Reuse algorithm presented, that is, the Hard Frequency Reuse, which consists of a segmentation of the total bandwidth and the assignment of a different sub-band to each cell. When this algorithm is applied, the ideal conditions of a zero TBLER, due to the highest level of SINR, and the maximum MCS index are satisfied. The expected results would be to reach the maximum throughput, since the condition are ideal. However, there is a limitation of the resources, which brings the scheduler to assign smaller TBs to each user, thus providing a smaller number of bits per second received by each one. Figure 4.1 depicts the curve of the average throughput of the UEs when Hard Frequency Reuse is adopted. It is demonstrated that, thanks to

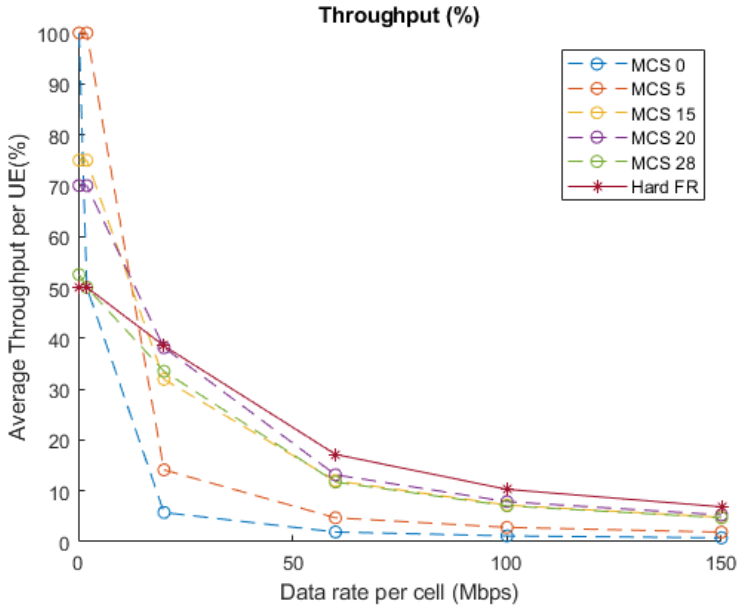

Figure 4.2: Average Throughput per UE (%) with different data rates per cell - NoOp and Hard Frequency Reuse algorithms.

a zero TBLER and the possibility to adopt a MCS index equal to 28 for each user, thus providing the highest spectral efficiency, the throughput remains restricted to low values in absolute terms. This is due to the fact that for each TTI there is the necessity to schedule many users as possible with, possibly, a large TBS. The reduction of the total bandwidth, due to the sub-band assigned to each cell, brings to a limited number of resources in terms of RBGs. In particular, as for a 20 MHz bandwidth the total number of RBs is equal to 106 and, as already said, the number of RBs grouped into a RBG is equal to 8, and since the sub-band reserves at most 17 RBs (that is, 106 RBs divided by 6 sectors), there are 2 assignable RBG within a TTI. Hence, because of the granularity affects the RBGs assignment, there is not the possibility to schedule more than 2 users per TTI. Furthermore, the Transport Block size designed is restricted by the small number of RBs available in the sub-band.

Moreover, figure 4.2 depicts the curves of the throughput in percentage values, intended as the reached value on the possible data rate achievable[2].

[2]The data rate reported on the x-axis in figure 4.2 is referred to the data rate provided by the cell. The bits per second received by a single UE within the cell can reach only that value divided by 5, since there are 5 UEs per cell. Then, for instance, if the cell data rate is 2 Mbps, and if all the UEs receive all the transmitted packets, they receive 400

Thus, it is worth to note that a lower data rate brings to a higher throughput even without Frequency Reuse algorithms since the spectral efficiency requested does not require the highest MCS index and, even with a lower SINR perceived by the users, all the packets are correctly decoded. Hence, if the data rate offered by the cell is about 200 kbps (with 40 kbps for each of the 5 UEs) all the users receive the total amount of transmitted packets even if the interference level is high. However, if a higher data rate is offered by the gNB, the necessary MCS index is different because of a higher spectral efficiency that must be reached in order to encapsulate many information bits as possible within a Transport Block. Thus, a Hard Frequency Reuse guarantees this efficiency thanks to the high SINR level and the consequent zero TBLER, so that the user is able to decode packets with a high modulation order, thus enhancing the ratio of received on the total transmitted packets. Figure 4.2 demonstrates that a higher bit rate requires a higher spectral efficiency, that is, a higher MCS index, which can be sustained with a hard reuse approach, that returns the best results in therms of throughput when a high data rate is offered by the cell. This value may increase by exploiting different Frequency Reuse algorithms, that optimize in a different way the use of the available resources.

Figure 4.1 reports also the throughput obtained when AMC is involved applying the NoOp algorithm. One may say that the MCS index selected for all the users in an adaptive way would be equal to MCS 20 because it is the value that guarantees the highest throughput between all the possible indexes. Figure 4.3 shows that the AMC provides a lower TBLER[3] compared to the case in which MCS 20 is selected for all the users. In fact, when MCS 20 is applied to all the users, the TBLER is equal to 0.75, while the Adaptive Modulation and Coding entity provides a TBLER equal to 0.35. This is due to the fact that the average MCS is not equal to 20 even if the throughput follows a trend similar to the violet curve of figure 4.1. Thus, the throughput is enhanced without increasing spectral efficiency for each user, but in an adaptive way, so that users which measure a high SINR level can sustain a high MCS index, while users far from the gNB are not able to decode packets with a high modulation order so that the MCS index value utilized for them will be lower, thus providing a lower TBLER thanks to the easier decoding process. By observing this behavior, it is worth to note that it may be possible to enhance the system performance by exploiting a higher

kbits per second, that is the 100% of the achievable bit rate, represented on the y-axis.

[3]The TBLER represented in figure 4.3 is computed by averaging the TBLERs obtained from the transmissions of packets to the single UEs within a specific TTI.

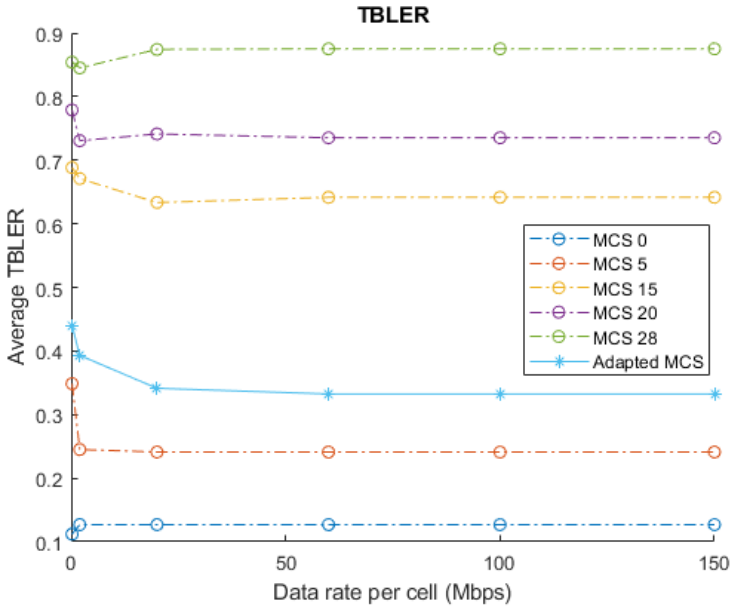

Figure 4.3: Average TBLER with different data rates per cell - NoOp algorithm.

MCS index for all the users, but in such a way that the TBLER level is kept low, thanks to a even higher received signal quality.

The two possibilities presented are the two basic ways to allocate resources so that, with the NoOp algorithm there is a higher interference level but a larger resources availability, while with a Hard Frequency Reuse the interference disappears but the effective assignable RBGs are less. `Ns-3` offers, as already mentioned, other Frequency Reuse algorithms, which are a mixture of these two methods and have the purpose to exploit all the resources in order to enhance the QoS offered to each UE.
By referring to figures 1.9 and 1.10, it can be seen how resources are assigned to UEs in different ways by classifying users into Center and Edge users depending on the RSRQ measured by each one. The Strict Frequency Reuse algorithm reserves edge sub-bands which do not interfere with the sub-band reserved to Center UEs. This may provide a higher SINR for both Center and Edge UEs but it limits the use of resources, as the single cell reserves both a sub-band for Center UEs and a sub-band for Edge UEs.
The Soft Frequency Reuse algorithm, in both its versions, is a mixture of NoOp and Hard algorithms: Center UEs exploit all the 20 MHz bandwidth (in the second version this does not apply, since the center sub-band is differ-

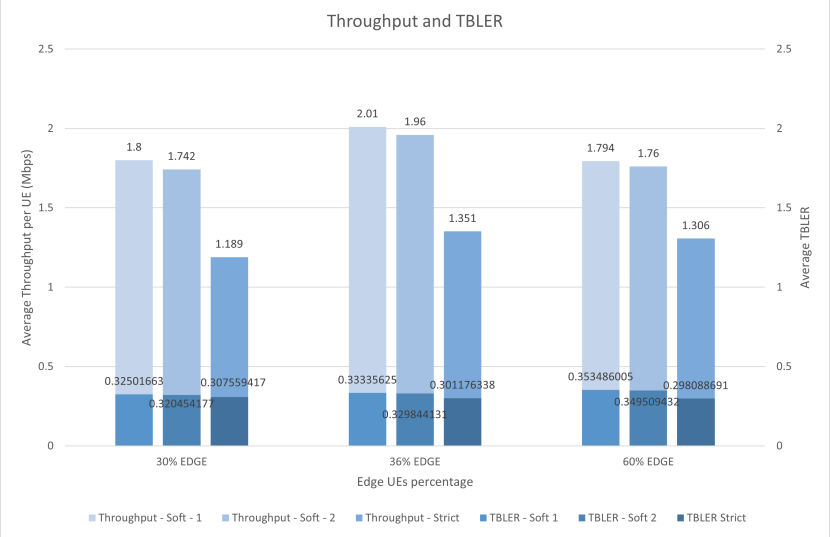

Figure 4.4: Average Throughput per UE (Mbps) for different Edge UEs percentages - Soft and Strict Frequency Reuse algorithms.

ent from the edge sub-band for each cell), while a different edge sub-band is reserved to Edge UEs in each cell. However, with respect to Strict Frequency Reuse, in this case the edge sub-bands overlap the total bandwidth reserved to Center UEs. This may imply a higher level of interference perceived by Center users of the neighbouring cells.

The simulation results in figure 4.4 are obtained with 5 UEs per cell and a data rate per cell equal to 150 Mbps. The graph reports the throughput calculated by averaging the throughput of all the users. Since UEs are classified into two categories based on the reported RSRQ, the number of Edge UEs depends on the adopted RSRQ threshold. Then, by increasing this threshold, the number of Edge UEs is higher due to the fact that the requested RSRQ is even higher. The setting of the RSRQ threshold is based on empirical results: by enhancing the threshold level, that is, when the 60% of users are Edge UEs, the throughput is lower than the optimum case, as well as for a lower threshold level, which counts more Center UEs.
The figure reports also the TBLER values obtained in the various simulations. As for the best case TBLER is similar to the one obtained when no ICIC is applied, the improved throughput is attributable to the fact that Soft Frequency Reuse exploits a higher spectral efficiency, and especially it provides a higher signal quality to those users that are located far from the gNB. In fact, by reserving edge sub-bands to those users, the signal quality

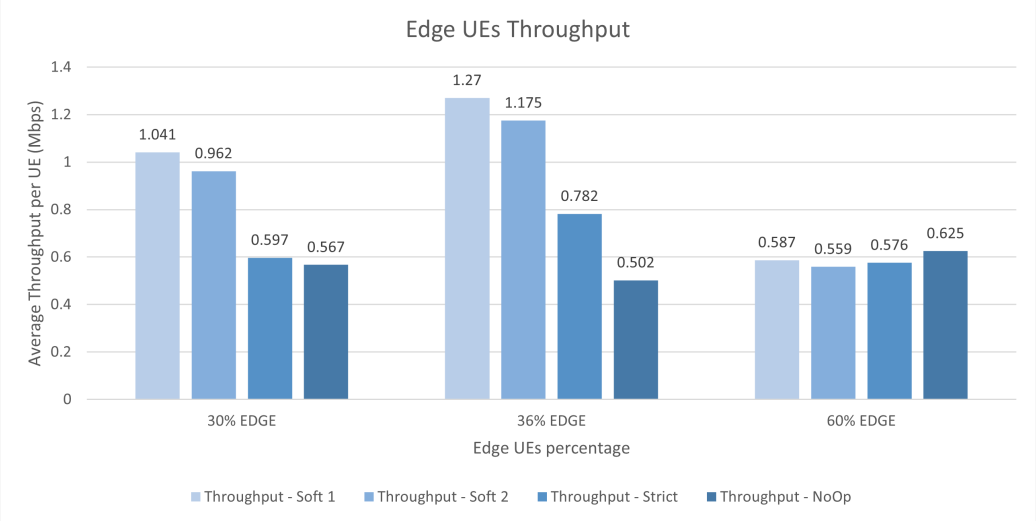

Figure 4.5: Average Throughput per UE (Mbps) for Edge UEs - Soft and Strict Frequency Reuse algorithms.

received by them is improved in such a way that TBLER is equal to zero even when a high MCS index is adopted for transmissions, thanks to the fact that the SINR level is higher. Then, the average throughput is enhanced by these users, which exploit a higher spectral efficiency. In particular, as depicted in figure 4.5, the throughput measured by Edge users is enhanced by Frequency Reuse algorithms. Edge users do not lose packets because of the high SINR, thus the average TBLER is increased only by Center UEs, which still perceive a certain interference level. In other words, by referring to the NoOp algorithm, since the TBLER computed is similar to the result provided by Soft Frequency Reuse, the improvement in terms of throughput is brought by those users that present a low throughput when NoOp algorithm is applied, which is enhanced when Soft Frequency Reuse is adopted, while Center UEs perform the same results in both the two cases. The fact that Center UEs still perceive a low SINR may be resolved by applying beamforming, whose results will be shown in the following.

Figure 4.5 compares the throughput computed by Edge users when three different Frequency Reuse algorithms are applied and when the NoOp algorithms is utilized. The users taken into account are those classified as Edge users by Frequency Reuse algorithms. The results show that all these users, which are distant from the gNB and may have a smaller number of decoded packets within a 1 s simulation due to the high interference level, can reach a higher throughput when Frequency Reuse algorithms are applied. The best

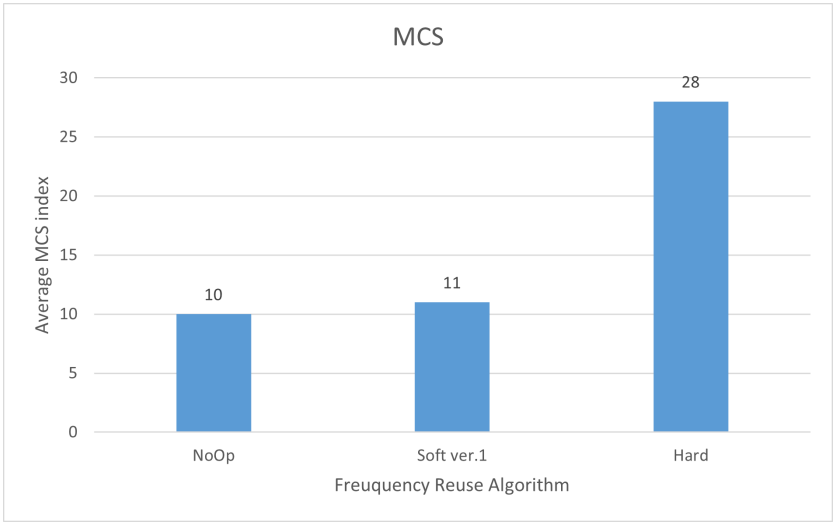

Figure 4.6: Average MCS index - NoOp, Hard and Soft Frequency Reuse algorithms.

result is provided by the Soft Frequency Reuse algorithm in its first version, with a RSRQ threshold which selects a number of Edge UEs equal to the 36%.

The throughput obtained by the simulations demonstrates that, among the three Frequency Reuse schemes just mentioned, the Soft Frequency Reuse (version 1) is the best scheme to adopt in this scenario. Since it provides an average throughput equal to 2.01 Mbps when the 36% of users are Edge users, it reaches a result, in terms of throughput per UE, which is similar to the one offered by the Hard Frequency Reuse. Thus, one may say that Soft Frequency Reuse algorithm is useful, as the best result is obtained with Hard Frequency Reuse, but it is worth to note that, in this case, the TBLER is not equal to zero and, as depicted in figure 4.6, the MCS index is not the highest. In fact, figure 4.6 represents the average MCS index for different adopted Frequency Reuse algorithms, which is calculated from the MCS indexes selected for transmissions to UEs. Then, whilst Hard Frequency Reuse provides the same results, in terms of throughput, as Soft Frequency Reuse, the latter is improvable. In fact, Hard Frequency Reuse already makes use of the highest MCS value and reaches the lowest TBLER, but it does not optimize the use of resources. Hence, Soft Frequency Reuse may be improved by applying beamforming, in order to increase the SINR measured by all the users such that TBLER can be reduced even when a higher MCS index is utilized.

4.2 Beamforming and Frequency Reuse algorithms

The results provided since here did not include a fundamental element of the 5G standard, that is, beamforming. Figure 4.7 shows the difference between the two basic methods already mentioned, with and without beamforming. The reported results refer always to the case in which a 150 Mbps data rate is offered by the single cell, in which 5 UEs are randomly located. The throughput calculated when no ICIC schemes are applied and only beamforming is adopted with all the UEs is even higher than the throughput returned by the simulations in which Hard Frequency Reuse is utilized together with beamforming. This is due to the fact that the only limit of the latter is the poor available resources when the total bandwidth is split in six parts, while the conditions are ideal, since each UE reports the highest SINR value and the TBLER is equal to zero. Then, starting from the NoOp algorithm, if beamforming is applied and also a Frequency Reuse algorithm such as Soft Frequency Reuse is adopted, the final result may enhance the system performance.

Then, the purpose is to exploit beamforming together with a method for which the available resources are allocated in such a way that a better throughput can be guaranteed to each user. Moreover, since the request is to increase the SINR measured by each UE in order to use a higher MCS

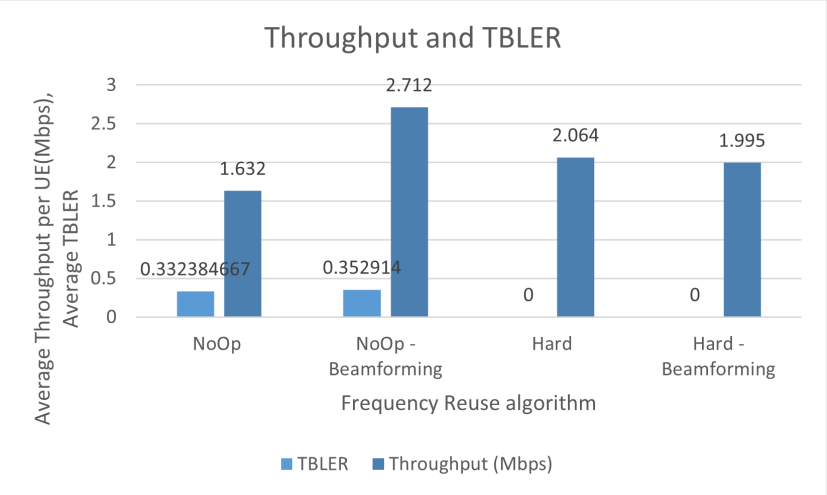

Figure 4.7: Average Throughput per UE (Mbps), and Average TBLER - NoOp and Hard Frequency Reuse algorithms with Beamforming.

index when transmitting data, so that each one can decode the received bits without packet loss, and also to exploit the maximum number of resources, the beamforming may be applied only to a certain number of UEs. If a Soft Frequency Reuse algorithm is utilized, there is a percentage of UEs that already measures a high SINR without beamforming. Therefore, it could be useless to direct beams toward a even higher number of UEs, as it may produce a higher interference level. In fact, if the power is directed toward a user located far from the gNB, it can result in a higher interference factor for neighbouring cells. Then, the possibility to direct beams only towards users that are located in the vicinity of the gNB may reduce the interference. The goal is to find out the optimum design for the two thresholds drawn in figure 3.10, as it is necessary to study the system behavior both when beamforming is applied to different numbers of UEs and when the number of Edge UEs is larger.

It is worth to remember that the choice of establish a beam from the gNB to the specific UE is made when the UE is attached to the network, while the classification of Edge and Center UEs is a subsequent choice. Then, the value of the distance threshold affects the design of the RSRQ threshold. In particular, if beamforming is used with a larger number of UEs, the interference that is generated over the neighbouring cells is greater than the interference produced when beamforming is utilized in a restricted area, since the beams are not directed towards the neighbouring cells. Figure 4.8 shows the percentage of Edge UEs present in the scenario when different thresholds values are set. As the distance threshold increases, the number of users for which beamforming is used is even larger. Thus, if the distance threshold is equal to 10 m and the RSRQ threshold is equal to 20 dB, the number of Edge users is higher than the case in which the distance threshold is larger. This is because the distance threshold must ensure a coverage at least to those users that are near the gNB. If this is not true, the number of users that perceive a low SINR level increases even when a low RSRQ threshold is set, since the requested RSRQ is low but there are more UEs which perceive a low SINR with respect to the case in which the distance threshold is larger. Furthermore, by increasing the RSRQ threshold it can be seen that the number of Edge UEs raises when the distance threshold increases. This is due to the higher interference level caused by the even larger distance reached by gNB beams, which interfere with the neighbouring cells. This requires a higher percentage of Edge UEs also because of the even higher RSRQ threshold.

In fact, when the RSRQ threshold is set to 30 dB and the distance threshold is such that a larger number of UEs is involved, the number of Edge UEs

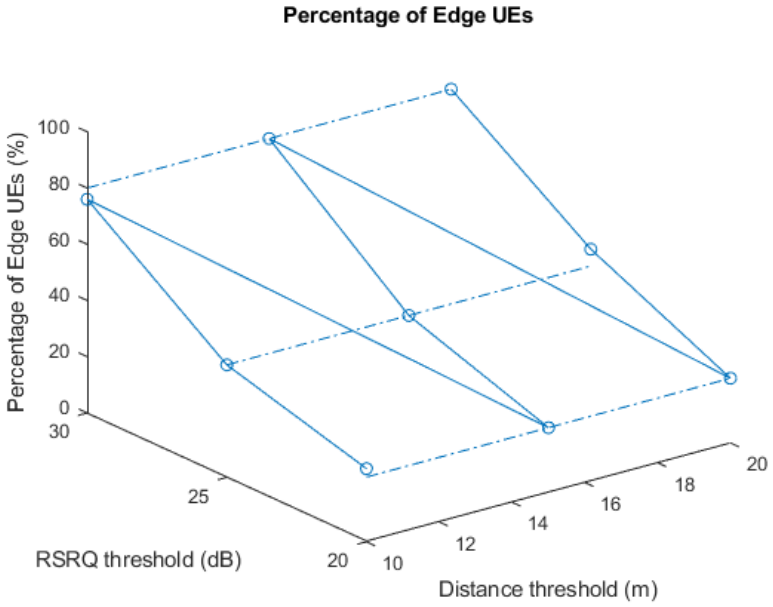

Figure 4.8: Percentage of Edge UEs with different Distance and RSRQ thresholds.

is equal to the 80%. Figure 4.8 may be utilized to evaluate the results by observing how the systems classifies users depending on the two thresholds.

Then, figure 4.9 shows the results obtained with different thresholds settings. It is clear that a high value for the RSRQ threshold does not provide a high throughput because of the tendency to a Hard Frequency Reuse behavior, which does not ensure the best performance. The greatest value in terms of throughput per user is obtained when a 25 dB RSRQ threshold is set with a distance threshold equal to 15 m. However, since fairness is a principle to keep in mind when such a scenario is studied, the best results in terms of throughput per Edge UE is obtained when the RSRQ threshold is equal to 25 dB and the distance threshold is 20 m. The risk is that, by expanding the area in which beamforming is applied, there could be the necessity to classify more users as Edge users, so that the approach would be tend to the Hard Frequency Reuse scheme and consequently the resources would not be exploited in the best way.

However, by evaluating the case in which beamforming is applied to all the users, the result is that shown in figure 4.10. This throughput value is obtained by directing beams towards each user, so that the interference

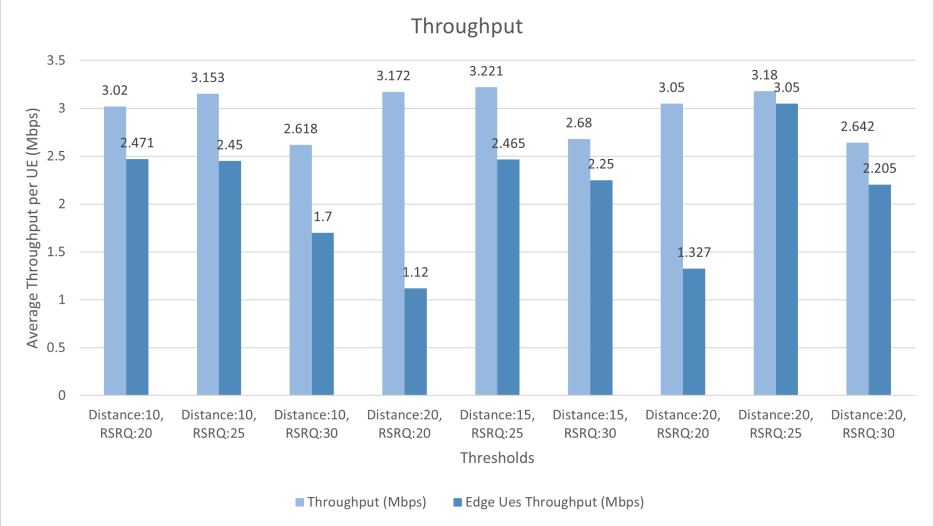

Figure 4.9: Average Throughput per UE (Mbps) for different Distance and RSRQ thresholds - Soft Frequency Reuse algorithm with Beamforming.

generated among neighbouring cells is higher, but it is also higher the signal power received by a user from its own cell. Thus, the fairness is ensured as the average throughput of all the users and the average throughput of edge users are similar. Also the number of Edge users does not increase significantly, as it corresponds to the 50% of all the users.

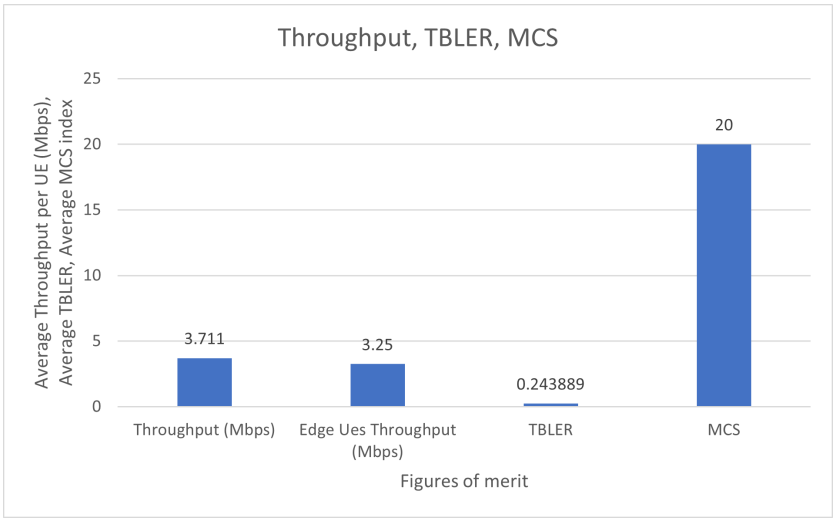

Figure 4.10: Figures of merit for the optimum RSRQ threshold - Soft Frequency Reuse algorithm with Beamforming.

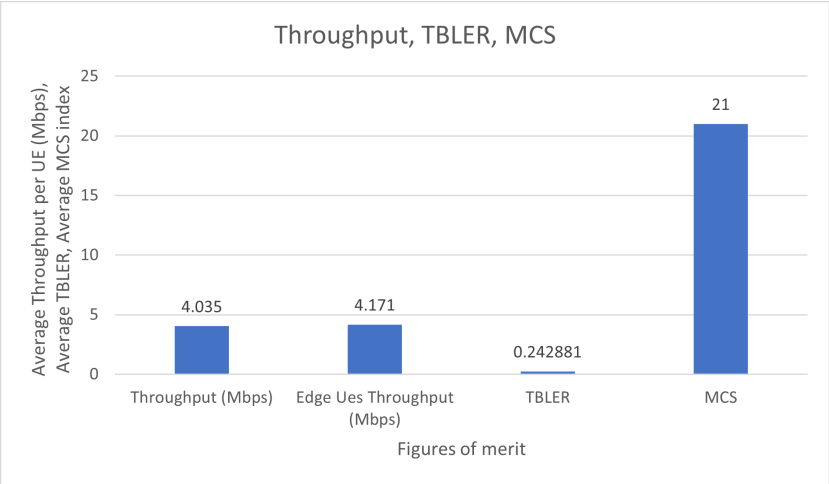

Figure 4.11: Figures of merit for the optimum RSRQ threshold - Soft Frequency Reuse algorithm with Beamforming - Proportional Fair scheduler.

Compared to the NoOp and the Hard Frequency Reuse algorithms, this configuration reports a MCS value equal to 20, which represents a high spectral efficiency, as the modulation order used for the allocated resources is equal to 6 and the code rate is equal to 0.554 (table 3.2), and it tends to the MCS 28 obtained with a Hard Frequency Reuse scheme. The TBLER, thanks to the fact that the SINR value measured by users is even higher, is reduced with respect to the NoOp algorithm, which returns a TBLER equal to 0.31, versus the 0.24 TBLER returned in this case.

However, the results proposed since here are provided by a Round Robin scheduler, which is probably the simplest scheduler found in the literature. It works by dividing the available resources among the active flows, i.e., those logical channels which have a non-empty queue. Thus, if a Proportional Fair scheduler is used, the prediction is that users which, at a given instant, perceive a better SINR will be preferred, in such a way that a higher spectral efficiency is achieved. On the other hand, the Proportional Fair scheduler ensures fairness if all the users can reach an equal SINR value in different instants, while it may fail when only a certain group of users can achieve appreciable SINR values. In other words, the Proportional Fair scheduler can be now utilized thanks to the fact that both Center and Edge UEs measure high SINR levels, while before it was not true, since there were not a method by which a better signal quality was ensured to each UE.
Figure 4.11 reports the results obtained with a Proportional Fair scheduler, where the RSRQ threshold is equal to 25 dB and beamforming is applied to

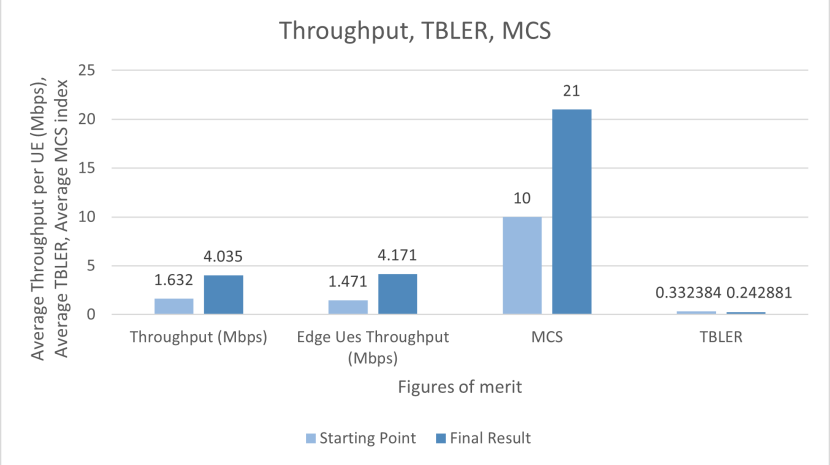

Figure 4.12: Comparison between the starting point and the final result.

each UE. The Soft Frequency Reuse (ver.1) shows that, with respect to figure 4.10, in which the same Frequency Reuse algorithm is used, this final result represents the best configuration obtained when both beamforming and Frequency Reuse algorithms are applied. The average MCS index value is equal to 21, which represents an important improvement in terms of spectral efficiency. Furthermore, the most significant aspect is that Edge UEs have a higher throughput than Center UEs. This is due to the fact that beams are directed towards all the users and, at the same time, Edge users already receive a high SINR level, thanks to Soft Frequency Reuse. The selection of such a scheduler gives priority to those UEs that perceive a high signal quality, that is, Edge UEs.

Hence, the final result is that a higher throughput can be reached by enhancing signal quality measured by the users, and this can be made by applying both beamforming and Frequency Reuse algorithms. Figure 4.12 represents the starting point, in which neither a ICIC scheme nor beamforming are adopted, and the final result, obtained with the combination of the already present feature of 5G, that is, beamforming, and an ICIC scheme, which is one of the possible Frequency Reuse algorithms and it has been implemented in `ns-3` to obtain these results. The figure shows that a higher throughput is received by users thanks to the exploitation of a high spectral efficiency, which is possible only if the SINR level perceived by each one is high. Then, if both the TBLER and the MCS are improved, it is possible to enhance the system performance. By referring to the starting point, in terms of throughput per UE, the improvement is about the 150%.

Conclusion

This work started from a 5G cellular scenario in which the interference was not coordinated, so that the resources were assigned to users favouring those which are located near the gNB and ignoring all those users that did not perceive a sufficient SINR level. Thus, the resources assignment were not optimized and, regardless of the total available bandwidth, users that were located far from the gNB were not served with sufficient information data. The application of Frequency Reuse algorithms together with the use of beamforming has been validated by `ns-3` simulations with the purpose to exploit all the bandwidth together with the aim of guarantee a higher spectral efficiency. Hence, by obtaining a high SINR level for all the users within the six different cells, the possibility to exploit a high modulation order and a high code rate brings to improved results.
However, the limitation of the total bandwidth makes the resources allocation difficult due to the fact that the scheduler must satisfy all the users within the cell but without disposing of all the necessary resources. Thus, the maximum value of the throughput can not be achieved.
The only application of Frequency Reuse algorithms limits the system performance and does not exploit an important 5G feature, which is beamforming. In fact, by directing a beam towards each user it is demonstrated, by observing the final throughput, that each user can reach a high SINR level: both Center and Edge UEs are served with the same priority and the achieved throughput represents a significant improvement from the starting point in which no ICIC scheme were applied. Moreover, the utilization of a Proportional Fair scheduler represents an improvement in terms of throughput thanks to the fact that it assigns resources to those users which perceive a higher SINR level, thus based on the instant channel quality of each user. Since each user can reach a high SINR level, thanks to both beamforming and Frequency Reuse algorithms, this type of scheduler allocates resources to users that instantly perceive the best channel quality, thus transmitting data with a higher spectral efficiency, which is the main parameter to increase when the purpose is to enhance the number of received bits per second.

This work demonstrates that, regardless of the 20 MHz bandwidth utilized, the resources are spread over all the users in the same way and, thanks to the applied algorithms, the starting situation is improved by 150%. Thus, if the initial configuration would bring the same results even with an infinite availability of frequency resources, the presented methods demonstrate that the total available bandwidth is exploited in such a way that all the users are served by the gNB with the same priority and the total throughput is enhanced thanks to the achieved average MCS which is even higher.

Finally, this work focuses on the application of ICIC schemes such as Frequency Reuse algorithms, but it does neither approach nor implement other possibilities for the reduction of interference, such as cell coordination-based schemes, which are dynamic ICIC schemes. Furthermore, this study does not even focus on interference mitigation schemes. Hence, all these applications may be investigated in future developments, starting from the results showed in this study, with the purpose to improve this work. Moreover, another possibility may be to exploit other 5G features, such as an increasing numerology or a transmission based on a FDD scheme, in order to verify if some improvements could be achieved.

List of Figures

List of Tables

Bibliography

[1] G. Fodor, C. Koutsimanis, A. Rácz, N. Reider, A. Simonsson, and W. Müller, "Intercell interference coordination in ofdma networks and in the 3gpp long term evolution system." *J. Commun.*, vol. 4, no. 7, pp. 445–453, 2009.

[2] A. S. Hamza, S. S. Khalifa, H. S. Hamza, and K. Elsayed, "A survey on inter-cell interference coordination techniques in OFDMA-based cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1642–1670, 2013.

[3] Ns-3 development team, *Ns-3 network simulator, ns-3 model library*, http://www.nsnam.org/docs/models/ns-3-model-library.pdf, March 2013.

[4] D. González, M. García-Lozano, S. Ruiz, and J. Olmos, "Static inter-cell interference coordination techniques for LTE networks: A fair performance assessment," in *International Workshop on Multiple Access Communications*. Springer, 2010, pp. 211–222.

[5] 3GPP, "5G; NR; Overall description; Stage 2," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.300, 01 2020, version 15.8.0.

[6] T. Zugno, M. Polese, N. Patriciello, B. Bojović, S. Lagen, and M. Zorzi, "Implementation of a spatial channel model for ns-3," in *Proceedings of the 2020 Workshop on ns-3*, 2020, pp. 49–56.

[7] 3GPP, "5G; Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.901, 05 2017, version 14.0.0.

[8] 3GPP, "5G; NR; Physical layer procedures for data," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.214, 10 2018, version 15.3.0.

[9] 3GPP, "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 36.331, 10 2018, version 15.3.0.

[10] 3GPP, "5G; NR; Multiplexing and channel coding," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.212, 07 2018, version 15.2.0.

[11] A. M. Cipriano, R. Visoz, and T. Salzer, "Calibration issues of PHY layer abstractions for wireless broadband systems," in *2008 IEEE 68th Vehicular Technology Conference*. IEEE, 2008, pp. 1–5.

[12] 3GPP, "Digital cellular telecommunications system (Phase 2+); Radio subsystem synchronization," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 45.010, 04 2011, version 10.1.0.

[13] P. Gawłowicz, N. Baldo, and M. Miozzo, "An extension of the ns-3 lte module to simulate fractional frequency reuse algorithms," in *Proceedings of the 2015 Workshop on ns-3*, 2015, pp. 98–105.

[14] N. Baldo and M. Miozzo, "Spectrum-aware channel and phy layer modeling for ns3," in *Proceedings of the Fourth International ICST Conference on Performance Evaluation Methodologies and Tools*, 2009, pp. 1–8.

[15] Z. Hanzaz and H. D. Schotten, "Analysis of effective SINR mapping models for MIMO OFDM in LTE system," in *2013 9th international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2013, pp. 1509–1515.

[16] C. Johnson, *5G New Radio in Bullets*. Independently Published, 2019.

[17] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2020.

[18] S. Sesia, I. Toufik, and M. Baker, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

Acknowledgements

Al termine di questa trattazione, ritengo di dover rivolgere i miei più sentiti ringraziamenti al professore relatore Ing. *Davide Dardari* e alla correlatrice Ing. *Marina Lotti*, per l'estrema disponibilità e per i preziosi insegnamenti e consigli che mi hanno impartito per la realizzazione di questo elaborato, facendomi appassionare ai temi trattati.
Desidero ringraziare il correlatore Ing. *Giovanni Ferri* che, insieme al team di JMA Wireless, mi ha accolto e mi ha aiutato con pazienza e grande disponibilità, fungendo da figura di riferimento e fonte di consigli per tutte le volte in cui ho avuto bisogno di un confronto.

Ringrazio tutta la mia *famiglia* per avermi dato la forza di affrontare con serenità questo percorso universitario; ringrazio i miei *nonni* per l'affetto e il supporto che mi hanno sempre dato, e mia *madre* e mio *padre* per avere permesso e incoraggiato la mia scelta di intraprendere gli studi in Ingegneria, ma soprattutto per avermi insegnato a coltivare con cura e dedizione le mie passioni e per avermi trasmesso la determinazione e la gioia nel cercare di superare sempre e comunque i propri limiti.

Un ringraziamento speciale va alla mia *Francesca*, che mi ha sempre sostenuto e ha sempre tirato fuori il meglio di me.

Infine ringrazio tutti coloro che mi hanno accompagnato in questa esperienza, per i momenti felici e di spensieratezza, per le emozioni, i pensieri e le esperienze che abbiamo condiviso insieme e che mi hanno arricchito umanamente, e per avermi donato il conforto e la leggerezza nei miei momenti più difficili.