PhD program in Network Engineering

Joint access-backhaul mechanisms in 5G cell-less architectures

Doctoral thesis by:

Rakibul Islam Rony

Thesis Advisors:

Dr. Elena López-Aguilera

Dr. Eduard Garcia-Villegas

Department of Network Engineering

Barcelona, Spain

March 2022

Acknowledgments

People say, doing a PhD is a solo journey. Fortunately for me, this saying proved to be wrong. During my journey, I have come across so many helpful people, genius minds, who have helped me to build the platform where I am standing today.

Firstly, I would like to express my utmost respect to my doctoral thesis supervisors, Dr. Elena López-Aguilera and Dr. Eduard Garcia-Villegas. They have helped me continuously throughout this journey not only to become a better researcher but also to grow as a better human being, I believe. I am very grateful that my supervisors have always made time for me to have weekly meetings. Additionally, they made me feel very welcome to discuss any issues and ask for help anytime I have encountered difficulties during this journey. During our numerous long discussions, they have always appreciated my ideas and with the help of their constructive comments, suggestions, and wisdom the ideas were translated into good research work. I strongly believe, this thesis itself and the related research publications are the results of excellent teamwork, where my supervisors played a very strong role.

My family has always been the strongest part of me. Their trust, support, and push have driven me towards achieving new milestones in life. I would like to take this opportunity to thank my mother Rasheda Sultana, my brothers Rashedul Islam Khan Rana, and Rabiul Islam Jony. Without their support, this would have not been possible.

My wife Jakiya Yesmin Eva is the one who actually took the stress similar to me during these years. She has supported me unconditionally during the downtimes, tolerated my excessive travel and workload, stood beside me with a smile during my little success stories. She took all the social, home, and family responsibilities so that I could totally concentrate on my work and finish my thesis to achieve my goal.

An honorable mention; Dr. Akshay Jain. He is my lab mate, my friend, my tutor, my guide, my well-wisher, my extra hand, my co-author. I have learned a lot from him about research, wireless technologies, optimization algorithms, coding, punctuality, work ethic, and kindness.

My gratitude to my supervisor during secondment, Dr. George Agapiou, and all the colleagues from OTE Academy, Greece, for welcoming me to their research lab. It was a

wonderful experience.

I would also like to thank UPC and EU Horizon 2020 research and innovation program, titled "Application-aware User-centric Programmable Architectures for 5G Multi-tenant Networks (5GAuRA)" for funding this research and giving me the opportunity to learn and grow as a researcher. I am also grateful to the colleagues and friends Mr. Matteo Vincenzi, Dr. Mikel Irazabal, Mr. Christos E. Tsirakis, Mr. Girma M. Yilma, Mr. Lanfranco Zanzi, Dr. Jian Song, Dr. Tareq Al-Shami, Dr. Mahmudul H. Kafi, Mr. Asif Habibi, Mr. Meysam Nasimi from 5GAuRA, who have taught me about teamwork. Thank you all for helping me and making me feel an essential part of the project deliverable, team meetings, training, summer and winter schools, workshops, open days, and many research events during this journey.

Finally, warm thanks to Google and YouTube :)

Dedicated to my father
Habibur Rahman Khan

This is for you Abba, we miss you

Abstract

Older generations of wireless networks, such as 1G and 2G were deployed using leased lines, copper, or fibre line as backhaul. Public Switched Telephone Network (PSTN) lines have also been considered as an option in a few cases. Voice traffic in 1G and 2G was simply supported by backhaul links, which evolved from a collection of Time Division Multiplexed (TDM) links. Later, in 2G and 3G, microwave wireless links have also worked as backhaul links while the backbone of the network was still wireline-based. However, due to multiple different use cases and deployment scenarios of 5G, a solo wireline-based backhaul network is not a cost-efficient option for the operators anymore. For cost-efficient and fast deployment, wireless backhaul options are very attractive. As for drawbacks, wireless backhaul links have capacity and distance limitations. To take advantage of both the aforementioned solutions, i.e., wired and wireless, 5G networks are anticipated to be deployed with heterogeneous backhaul networks. Previously, radio access designs considered the backhaul networks to be sufficient, which is certainly not the scenario in 5G networks due to their high capacity requirements. Therefore, the transport network of 5G is expected to be a costly component, because of its heterogeneity, complexity, and stringent requirements.

To address the aforementioned challenges, wireless backhaul options are providing more attractive solutions (i.e., cost-efficiency, fast and easy deployment), and hence, technologies using the same resources (e.g., frequency channels) may be used by both access and backhaul networks. In this scenario, resource sharing and cooperation between access and backhaul network can minimize the network deployment and maintenance cost significantly. To perform this cooperative operation in a cost-efficient way, blurring the separation line between access and backhaul networks becomes necessary. Therefore, in 5G the access and backhaul networks cannot be seen as separate entities; rather, integrated together to ensure the best use of resources, while independent optimization of the access network is not enough anymore.

Therefore, in this thesis, we provide a detailed discussion on joint access-backhaul mechanisms, a key enabler for 5G wireless networks. To do so, firstly in Chapter 1, we explain the complexity and heterogeneity of 5G transport network and its strict performance require-

ments. We also discuss available and future potential backhaul technologies considered to take the challenges.

In Chapter 2, a detailed state-of-the-art in joint access-backhaul mechanisms is presented, and different approaches to perform joint access-backhaul mechanisms are classified into: i) joint functional design, and ii) joint optimization.

In Chapter 3, an initial performance evaluation of access-aware backhaul optimization is presented, where the backhaul network is dynamically assigned with the required resources to serve the dynamic requirements of the 5G access network. The evaluation results and discussions manifest the resource efficiency of joint access-backhaul mechanisms.

Functional splits in different layers of the access network come as an intelligent solution to reduce the enormous capacity requirements of the transport network in a centralized radio access network approach, which tends to centralize almost all the functionalities into a central unit, leaving only radio frequency functions at the access points. In Chapter 4, we provide a detailed discussion on functional splits at the physical layer and their economic impact on the transport network design. From a joint access-backhaul functional design perspective, in Chapter 4, we propose a novel technique, which takes the benefit of functional splits at the physical layer to design a heterogeneous transport network in an economical budget and capacity-limited scenario.

As motioned earlier, wireless backhaul options are getting more attention due to their economic and ease of deployment benefits. However, the limited capacity of the wireless links remains a challenge, and subsequently, frequency spectrum becomes scarce, and therefore, requires efficient utilization. In Chapter 5, a joint spectrum sharing technique to implement joint access-backhaul mechanism is presented. Evaluation results show that our proposed joint spectrum sharing technique, where spectrum allocation in the backhaul network follows the access network's traffic load, is fair and efficient in terms of spectrum utilization. In Chapter 5, we also propose a machine learning technique, which analyses data from a real network and estimates the access network's traffic pattern, and subsequently, assigns bandwidth in the access network according to the traffic estimations. Presented evaluation results show that a well-trained machine learning model can be very efficient in predicting the network requirements and bandwidth allocation and obtain an efficient utilization of scarce resources, i.e., frequency spectrum.

In summary, in this thesis, we discuss the importance of joint access-backhaul mechanisms to achieve the performance goal of 5G. Our proposed novel techniques and the evaluation results show that such joint mechanisms can pave the way towards programmable, scalable, and resource-efficient 5G.

List of Publications

[J2] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Dynamic spectrum allocation following machine learning-based traffic predictions in 5G" IEEE Access 9 (2021): 143458-143472. (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)). doi: 10.1109/ACCESS.2021.3122331.

[J1] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Cost Analysis of 5G Fronthaul Networks Through Functional Splits at the PHY Layer in a Capacity and Cost Limited Scenario." IEEE Access 9 (2021): 8733-8750. (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)). doi: 10.1109/ACCESS.2021.3049636.

[C4] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Optimization of 5G fronthaul based on functional splitting at PHY layer." In 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1-7. IEEE, 2018. doi: 10.1109/GLOCOM.2018.8647928.

[C3] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Cooperative spectrum sharing in 5G access and backhaul networks." In 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 239-246. IEEE, 2018. doi: 10.1109/WiMOB.2018.8589187.

[C2] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Access-aware backhaul optimization in 5G." In Proceedings of the 16th ACM International Symposium on Mobility Management and Wireless Access, pp. 124-127. 2018. doi: 10.1145/3265863.3265881.

[C1] R. I. Rony, A. Jain, E. Lopez-Aguilera, E. Garcia-Villegas, and I. Demirkol, "Joint access-backhaul perspective on mobility management in 5G networks", IEEE Conference on Standards for Communications and Networking (CSCN) 2017, pp. 115–120, Sept. 2017. doi: 10.1109/CSCN.2017.8088608.

Index

List of Figures

List of Tables

List of Abbreviations

2G	Second generation
3G	Third generation
3GPP	Third Generation Partnership Project
4G	Fourth generation
5G	Fifth generation
5G NORMA	5G Novel Radio Multiservice adaptive network Architecture
5G NR	5G New Radio
5GPPP	The Fifth Generation infrastructure Public Private Partnership
ADSL	Asymmetric Digital Subscriber Line
AI	Artificial Intelligence
ANN	Artificial Neural Networks
AP	Access Point
ABS	Almost Blank Sub-frame
AD	Analog to Digital
B-PON	Broadband Passive Optical Network
BBU	Baseband Unit
BCUF	Backhaul Capacity Utilization Factor
BER	Bit Error Rate
BH	Backhaul
BR	Bayesian Regularization
BRAN	Broadband Radio Access Networks
BS	Base Station
BSC	Base Station Controller
BW	Bandwidth
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CDF	Cumulative Distribution Function
CF	Carrier frequency

CGMII	100 Gbps Media Independent Interface
CN	Core Network
CoMP	Coordinated Multi-Point transmission
CP	Control Plane
CRAN	Centralized Radio Access Network
CRE	Cell Range Extensions
CREO	Cell Range Extensions offset
CU	Centralized Unit
CWDM	Coarse Wavelength Division Multiplexing
D2D	Device-to-Device
DL	Downlink
DP	Data Plane
DRAN	Distributed Radio Access Network
DSA	Dynamic Spectrum Allocation
DU	Distributed Unit
DWDM	Dense Wavelength Division Multiplexing
E2E	End-to-end
EDGE	Enhanced Data Rate for GSM Evolution
eMBB	EEnhanced Mobile Broadband
eNB	Evolved Node-B
eICIC	Enhanced Inter-Cell Interference Coordination
EPC	Evolved Packet Core
EPON	Ethernet PON Passive Optical Network
ETSI	European Telecommunications Standards Institute
E-PON	Ethernet Passive Optical Network
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FCUF	Fronthaul Capacity Utilization Factor
FD	Full Duplex
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FH	Fronthaul
FLEX-RAN	Flexible Radio Access Network
FSO	Free Space Optical
FP7	7th Framework Programme for Research and Technological Development
FR	Frequency Reuse
FSA	Fixed Spectrum Allocation

FSO	Free Space Optics
FTTH	Fibre To The Home
G-PON	Gigabit Passive Optical Network
GA	Genetic Algorithm
GB	Gigabyte
GF	Gradient Follower
gNB	next generation NodeB
GPRS	General Packet Radio Services
GSM	Global System for Mobile Communications
HARQ	Hybrid Automatic Repeat Request
HD	Half Duplex
HetNet	Heterogeneous Networks
HSPA	High Speed Packet Access
IBFD	In Band Full Duplex
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunications
IoT	Internet of Things
ISD	Inter Site Distance
ITU-R	International Telecommunication Union-Radio Communication Section
ITU-T	International Telecommunication Union-Telecommunication Standardization Sector
JT-CoMP	Joint Transmission CoMP
KPI	Key Performance Indicator
LM	Levenberg Marquardt
LOS	Line of Sight
LSTM	Long-Short Term Memory
LTE	Long Term Evolution
LTE-A	Long Term Evolution Advanced
MAC	Medium Access Control
MBS	Macro-Base Station
MC	Macro-cell
MADM	Multiple Attribute Decision Making
MH	Midhaul
MIMO	Multiple Input and Multiple Output
MiWEBA	Millimetre-Wave Evolution for Backhaul and Access
ML	Machine Learning
MM	Mobilty Management

mmWave	Millimetre Wave
MN	Mobile Node
MSE	Mean Squared Error
MN	Mobile Node
MTC	Machine Type Communication
NARNET	Nonlinear Autoregressive Neural Network
NFV	Network Function Virtualization
NG-PON	Next Generation Passive Optical Network
NGC	Next Generation Core
NGFI	Next Generation Fronthaul Interface
NGPON	Next Generation Passive Optical Network
NLoS	Non-Line of Sight
NN	Neural Networks
OADM	Optical Add Drop Multiplexing
OBFD	Out-of-band-full-duplex
OF	OpenFlow
OBSAI	Open Base Station Architecture
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLT	Optical Line Termination
ONU	Optical Network Unit
OPEX	Operating Expenditure
PDCP	Packet Data Convergence Protocol
PHY	Physical
PON	Passive Optical Network
PtMP	Point-to-Multi-Point
PtP	Point-to-Pointt
QoE	Quality of Experience
QoS	Quality of Service
RANaaS	Radio Access Network as a Service
RAT	Radio Access Technology
RE	Resource Elements
RF	Radio Frequency
RLC	Radio Link Control
RNC	Radio Network Controller
RNN	Recurrent Neural Networks

RoF	Radio-over-Fiber
RPR	Resilient Packet Rin
RRC	Radio Resource Control
RRH	Remote Radio Head
RRM	Radio Resource Management
RRU	Remote Radio Unit
RSRP	Reference Signal Received Power
RSSI	Received Signal Strength Indicator
SC	Small-cell
SCF	Small Cell Forum
SCG	Scaled Conjugate Gradient
SDM	Space Division Multiplexing
SDN	Software Defined Networking
SE	Spectral Efficiency
SF	Scalig Factor
SI	Self Interference
SINR	Signal to Interference and Noise Ratio
SLA	Service Level Agreement
SSA	Spectrum Sharig Approach
SU	Secondary Users
SVM	Support Vector Machines
TCO	Total Cost of Ownership
TDD	Time Division Duplex
TDM-PON	Time Division Multiplexing Passive Optical Network
THz	Teraheartz
UCB	User-Centric Backhaul
UDN	Ultra-Dense Network
UE	User Equipment
UF	Utilizatio Factor
UL	Uplink
UMTS	Universal Mobile Telecommunications Systems
V2X	Vehicle to everything
VDSL	Very high-speed Digital Subscriber Line
VNF	Virtualized Network Function
VR	Virtual Reality
WBH	Wireless Backhaul Hub

WDM-PON Wavelength Division Multiplexing Passive Optical Network

Chapter 1

Introduction

To meet the Key Performance Indicators (KPI) promised by subsequent generations of wireless communications, Radio Access Networks (RAN) had to evolve with different functional and architectural changes. A basic cellular network consists of User Equipment (UE), Base Stations (BS), and Core Network (CN). Through the years of evolution, the functionalities and capabilities of these blocks have been changing. At the beginning, in 1st Generation (1G) and 2nd Generation (2G), BS performed most of the RAN functionalities. Each BS was connected to the CN via a backhaul network. These traditional BSs performed the signal processing, transmission, and reception with the UEs. Several BSs formed a cluster, which is controlled by Base Station Controller (BSC). The BSs were connected to each other to perform handovers for moving UEs. These BSs were totally equipped and placed in rooftops or wall mounted with a few 10m of height. Later, 3rd Generation (3G) network architecture distributed the functionalities of BSs in Universal Mobile Telecommunications System (UMTS) architecture. Radio Network Controller (RNC) was introduced to control the resources and functionalities centrally [4]. 3G system supported different types of BSs with different ranges of coverage area and different functionalities. Hence, the cost of deployment is reduced and the coverage is increased. Small cells (SC) (i.e. femto-cells, pico-cells, metro-cells) came into the network as low cost, low power, energy-efficient Access Points (APs). These small cells/micro-BSs are connected to the BSs/macro-BSs via a wireless or wired backhaul. Later, the Evolved UMTS Terrestrial Radio Access Network (EUTRAN) introduced Evolved Node-B (eNB) replacing BSs, and the functionalities centralized in RNC were distributed to the eNBs [4]. In this Distributed RAN (DRAN), eNBs are connected to each other via the X2 interface and to the CN via the S1 interface. Another approach, Centralized RAN (CRAN) proposes full centralization of functions, employing only RRH (Remote Radio Head) in the network placed very close to the UEs. RRH performs only the

Radio Frequency (RF) functionalities and all other functionalities are centralized in Base Band Unit (BBU). In future networks, different cells with different coverage areas, purposes, and RATs will form Heterogeneous Networks (HetNets) in the access network. Additionally, macro-BSs (MBS) will ensure the coverage, while small cells will extend the coverage and boost capacity where required. Some of the MBSs can include more functionalities such as a backhaul aggregator for the small cells connected to it, or the backhaul aggregator can exist as a separate entity. The small cells can be connected to the BBU directly or via MBSs or backhaul aggregators using wired/wireless backhaul links. Figure 1.1 gives a brief view of RAN evolution, which has been discussed so far.

As mentioned, the capacity demand of mobile networks is exponentially increasing throughout the years with the evolution of RAN. Every decade, a new generation of RAN arrives with new features and subsequently boosts up the expectations. With the increased demand for data-rich applications (video call, online streaming) and data-rich devices (smartphones, tablets, etc.) wireless communications system is expected to provide around one hundred times more capacity (i.e., 10 Mbps per m^2) to a hundred times more connected devices (i.e., 10^6 connected devices per km^2) compared to 4G. Additionally, the expected end-to-end latency for delay-sensitive applications is less than 1ms [C1]. According to [5], in 2020, the number of mobile subscriptions has totaled 7.9 billion globally, and mobile data traffic grew by 49% within the previous year. By the end of 2026, the amount of mobile subscriptions is expected to be 8.8 billion, among which, 88% will consist of mobile broadband connections. Furthermore, with this mobile broadband connectivity, each user will be expecting a huge capacity, i.e., 25 Gigabyte (GB) on average per month [5]. Therefore, the Fifth Generation (5G) of cellular networks is here with a vision to meet the future demands in terms of capacity, coverage, and Quality of Service (QoS). To do so, 5G already identified a few key enablers such as CRAN, Network Function Virtualization (NFV), Software Defined Network (SDN), Multi-tenancy, Multi-Radio Access Technologies (RAT), functional split, network slicing, mmWave, Massive-Multiple Input Multiple Output (MIMO), Device-to-Device (D2D) communications, massive deployment of Internet of Things (IoT), etc [6]. With the aforementioned enablers, 5G is promising to provide ubiquitous coverage and very high data rates in the access networks.

From the operators' point of view, infrastructure sharing is already a popular solution as cost efficient deployment. Moreover, the 5G network will allow resource and network sharing among operators. Network slicing will enable the sharing of the same mobile infrastructure between different operators while having their own logical network [7]. Additionally, multi-tenancy will ensure the best use of available resource sharing including infrastructure and spectrum. With all these factors in mind, the RAN architecture of 5G has to be more flexible,

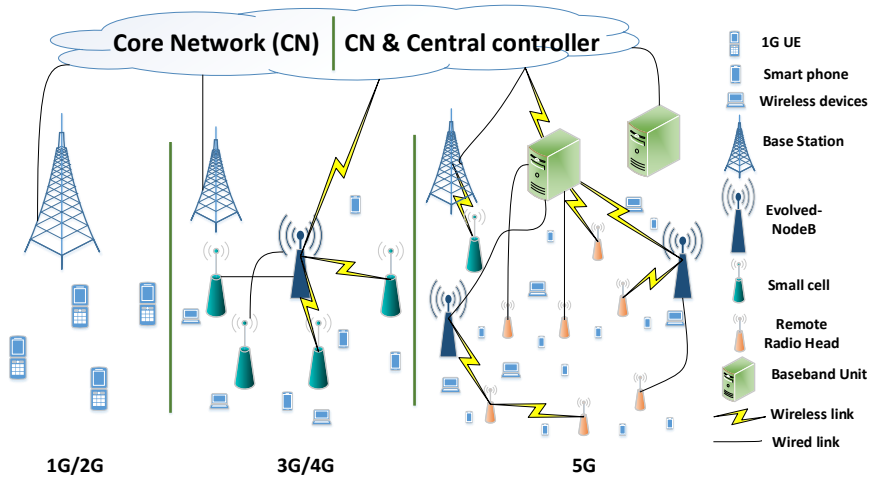

Figure 1.1: Evolution of Radio Access Network.

scalable, energy-efficient, and cost-effective.

Nonetheless, the aforementioned disruptive features of 5G to deal with future demands are making the access networks more complex, and require a unique backhaul network to carry the large amount of data in a very short time. In other words, future access networks have to be supported by a high capacity, low latency transport network. According to [8], 5G RAN will require a backhaul capacity at least 10 times than what is demanded by 4G RAN. In this scenario, designing such a transport network becomes challenging, and, as discussed in the following, no single backhaul solution can meet the expected requirements. Thus, it 5G backhaul network must be heterogeneous, employing different solutions according to corresponding individual requirements.

Moreover, CRAN is considered a key enabler of 5G. CRAN includes high requirements in the backhaul network, due to the fact that proper communication between RRH and BBU, and between BBU and core, are needed, and a large amount of data needs to be carried in a very small amount of time. In this way, the performance of 5G RAN/CRAN is highly dependent on the backhaul capacity and latency. In this context, joint optimization of access network along with backhaul network is necessary, where access and backhaul can influence each other [9]. In addition, 5G systems should allow multi-RAT and multi-connectivity in the network, where different RATs can serve simultaneously the user according to its requirements. Besides, the cell-less architecture concept, where all the APs are coordinated with each other, allows the user to employ resources from different APs and RATs in order to get the best performance. All of these new concepts require a unique backhaul, which is fully aware of the access network. To validate these features, access and backhaul have

to work jointly, while the optimization can be dependent on each other. In other words, backhaul resources should be assigned according to the requirement of access network, i.e., some APs having more traffic to serve with higher data rates can be assigned with more backhaul resources, and they can be taken away while no longer required.

Focusing on the required capacity of 5G backhaul networks; to be able to provide such large capacity (i.e., 10 times compared to 4G), the backhaul resources (i.e., bandwidth, available link) cannot be considered overprovisioned anymore. Hence, the adaptation of newer technologies (i.e., mmWave) to provide higher capacity and validation of efficient utilization of the available resources will be unavoidable. On the other hand, to meet the RAN requirements, 5G will densely deploy mmWave based SCs to enhance the capacity in the access networks as well. In this scenario, both the access and backhaul networks may employ the same technology (i.e., mmWave), where separating both the entities will not be efficient [10], rather; the line between the access and backhaul networks will be blurred. Moreover, in future dense networks, UEs will be able to receive signals from several access points, and proper communication among APs to ensure the QoS is required. To validate the aforementioned idea, the concept of "joint access-backhaul operation" comes in, which is considered as an additional key enabler of 5G. In joint access-backhaul operation, both networks work closely with each other while being aware of each other's current situation and requirements. To be able to adapt to each other's conditions, both the networks can optimize themselves and make sure the best usage of available resources efficiently. In this way, joint access-backhaul operation minimizes the Capital Expenditure (CAPEX) and Operational expenditure (OPEX) [C1], enabling network programmability, flexibility, elasticity, and efficiency.

The aforementioned joint access-backhaul mechanism can be categorized as, i) joint functional design, where both networks are functionally integrated to relax the stringent requirements over the backhaul network, and, ii) joint optimization of access and backhaul networks, where access and backhaul networks are optimized together depending on each other's requirements and current conditions. Both approaches have their own challenges and opportunities. In this research, we take both of them into account to study, propose and evaluate joint access-backhaul mechanism in 5G networks.

1.1 Research motivation

Joint access-backhaul mechanisms create new opportunities in the design of 5G networks. As discussed earlier, this paradigm is unavoidable to ensure the best usage of precious resources for both access and backhaul networks, which makes 5G networks more cost and resource-

efficient. However, as discussed in the subsequent chapter, there can be many different areas to perform joint optimization. In this research, we focus on the following challenges to perform joint optimization of access and backhaul networks.

- 5G networks are expected to have ultra-densely deployed cells. Moreover, the huge amount of devices connected to these cells requires a very large backhaul capacity. Therefore, it is becoming more challenging to design such high-capacity backhaul networks. Additionally, D2D communication, Machine Type Communications (MTC), and IoT services require very low latency in the transport network which further limits the backhaul technology options.

- Joint operation of access-backhaul networks requires higher flexibility in the network, where resources are expected to be allocated dynamically and on-demand basis. Additionally, resource allocation schemes in backhaul and access networks will have to be aware of each other's conditions and requirements. Subsequently, new resource allocation and optimization schemes are required to be realistic and simple for validating the joint access-backhaul mechanisms.

- With integrated RAN-backhaul design, different functional splits at the Physical (PHY) layer can be utilized to relax the enormous capacity requirements on the transport network, however, at the cost of loosing the centralization gain. Moreover, it has a great impact on CAPEX and OPEX of the network. Additionally, this mechanism makes the design of RAN more complex. Therefore, finding the optimal functional split at the PHY layer for different APs according to the available budget and the available backhaul resources can be challenging.

- The limited available spectrum is one of the most challenging issues when designing wireless networks, where a massive amount of users will be asking for huge data rates. Studies show that re-use of the available spectrum is the most effective and cost-efficient way to increase the network capacity. However, re-use of spectrum imposes interference in the network. Thus, efficient spectrum management techniques are required. From this research point of view, backhaul and access links sharing the same spectrum in two-tier or multi-tier networks will make the networks more complex and challenging to design. And thus, a proper joint access-backhaul spectrum sharing technique is required to ensure the best use of spectrum.

1.2 Research Objectives and Contributions

Challenges come often with multiple new research opportunities. The shortcomings of the backhaul options in terms of capacity can be addressed by their developments, which might provide more suitable options for 5G. mmWave is already a popular option but proper models are yet to be developed. Moreover, even higher frequencies, such as the Terahertz (THz) band, can be taken into account as they offer a great amount of spectrum resources, which can provide a data rate of hundreds of Gbps [6].

As for the huge capacity requirements imposed on the transport networks, various data compression techniques can be adopted for relaxing the backhaul requirements. In this regard, a trade-off between loss and gain due to compression needs to be studied. Functional splits in Medium Access Control (MAC) and Radio Resource Control (RRC) layer further relax those backhaul requirements. Besides, the necessity for standardization provides the opportunity for more focused research directions.

In the following, we list the objectives together with the contributions of this research focusing on joint optimization of access and backhaul networks in 5G, to tackle the challenges listed in Section 1.1:

- *State-of-the-art study*: Joint access-backhaul mechanism is a very recent paradigm for wireless networks. Therefore, the first task of this research effort is to develop the state-of-the-art in order to identify related works and their advancements. Subsequently, the open issues and challenges on the topic are more visible. After a thorough literature review, a detailed discussion on different approaches to perform joint access-backhaul mechanism was presented in our publication, titled "Joint access-backhaul perspective on mobility management in 5G networks" [C1], which was presented at the *2017 IEEE CSCN* conference. Additionally, in this work, future complex transport network composed of fronthaul, the connecting link between BBU and RRH; midhaul, the link connecting BBUs to an aggregated fronthaul point; and backhaul, the link connecting BBU to the core network, were thoroughly discussed. Finally, we have also provided a discussion on how the Mobility Management (MM) in 5G can be a part of such joint access-backhaul optimization.

- *Access-aware backhaul optimization*: In future networks, backhaul resources can not be considered overprovisioned anymore. Hence, one of the most important objectives of this research is to develop and propose techniques to ensure the best usage of scarce resources and contribute to the cell-less concept. Therefore, contrary to traditional approaches, we propose more concerned backhaul resource distribution techniques, which

work on an on-demand basis. Doing so, our paper, titled "Access-aware backhaul optimization in 5G" [C2] proposes a resource-efficient access-aware backhaul optimization technique, where backhaul resources are optimized on-demand according to the current requirements of access networks. This work was presented at the *2017 ACM MobiWAC* conference.

- *Functional splits at the PHY* layer add another degree of flexibility to adapt the CRAN with the available fronthaul resources. On the other hand, different splits at the PHY layer transform into different levels of capacity requirements on the fronthaul link and, subsequently, affects CAPEX and OPEX of the network. Therefore, an optimal split at the PHY layer to deal with the fronthaul capacity requirements and to perform an integrated access-backhaul design is required. Our paper, titled "Optimization of 5G fronthaul based on functional splitting at PHY layer" [C4] proposes and evaluates a unique optimization of PHY layer splits in an OPEX and capacity-limited scenario. This work was presented at the *2018 IEEE GLOBECOM* conference. That work was followed by a Total Cost of Ownership (TCO)-based optimization of PHY layer splits. To do so, a discussion on the feasible technologies for fronthaul network and the respective cost analysis is also required. Our paper, titled, "Cost Analysis of 5G Fronthaul Networks Through Functional Splits at the PHY Layer in a Capacity and Cost Limited Scenario" [J1] presents a detailed discussion on the feasible wireless and wired technologies for transport networks (fronthaul link) varying with the PHY layer splits and the subsequent requirements on the transport network. This work was published in the IEEE Access Journal (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)).

- *Joint spectrum sharing*: To be resource and cost-efficient, in future networks the scarce spectrum will be shared between access and wireless backhaul networks from a common spectrum pool, where the entire available spectrum is gathered. Subsequently, cooperative access-backhaul mechanisms become necessary to ensure the best use of the valuable resources, i.e., bandwidth. Hence, in our paper, titled "Cooperative spectrum sharing in 5G access and backhaul networks" [C3], we present and evaluate the idea of spectrum sharing among different links from a joint access-backhaul mechanism point of view. This work was presented at the *2018 IEEE WiMOB* conference. The presented evaluation shows that a proper cooperative spectrum sharing between access and backhaul networks can result in a better-performing network in terms of spectral efficiency and resource distribution fairness. This work was followed by an extensive work, titled, "Dynamic spectrum allocation following machine learning-based traffic

predictions in 5G" [J2]. This work is published in the IEEE Access Journal (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)). In this work, we have proposed and investigated a dynamic spectrum allocation approach among APs from a common spectrum pool, according to a cell-less architecture concept, which follows real network requirements expressed through the operator's network traces. We have also proposed a unique technique to predict the access network requirements utilizing machine learning, and showed how spectrum can be accurately allocated following the dynamic predictions.

1.3 Thesis Outline

After discussing the research challenges, objectives and contributions, in this section we provide an outline of the upcoming sections of this thesis.

Our contributions to this research work are divided into four different areas. Firstly, Chapter 2 discusses the state-of-the-art of joint access-backhaul mechanisms. To do so, we first focus on the future complex architecture of the backhaul network and, subsequently, a detailed study on potential backhaul technologies is presented. Afterward, we go through different research projects and their prospects on join access-backhaul mechanisms. Finally, different approaches to perform joint access-backhaul are categorized into, joint access-backhaul functional design and joint access-backhaul optimization.

Secondly, in Chapter 3, access-aware backhaul optimization is presented, where BH resources are offered to the BH links according to their current requirements, which follows their respective access network conditions. This mechanism also allows saving the scarce resources, and re-using them accordingly. In this section, we present and evaluate the stated idea to perform joint optimization of access ad backhaul resources according to their current requirements.

Thirdly, in Chapter 4, a detailed discussion on functional splits at the PHY layer, related cost distribution, OPEX, and TCO-based optimization of PHY layer splits are presented. This chapter shows the potential of functional splits at the PHY layer to design a frothaul network in a cost and capacity-limited scenario.

Finally, in Chapter 5, contributions for spectrum sharing in future networks are presented, where access and BH networks are expected to share a common spectrum pool. The proposed approach consists in a joint access-backhaul awareness mechanism; we propose a spectrum allocation mechanism, which considers the current access and backhaul networks requirements. The spectrum allocation approach in a dynamic access network is later discussed in the same chapter. Machine Learning-based traffic prediction, and the potential benefit

of dynamic access network's requirement-aware spectrum allocation, are also presented and supported with evaluation results in this chapter.

Chapter 6 concludes the thesis and presents future potential research tracks of the presented contributions.

Chapter 2

State of the Art in Joint Access-Backhaul mechanism

This chapter highlights and summarizes the existing related works on the joint access-backhaul mechanism. Firstly, we mention the related projects with a focus on this paradigm and their respective contributions/objectives. Later, we categorize joint access-backhaul mechanisms into two different approaches and corresponding works are discussed accordingly. However, before diving into the related work, it is important to look deeper into the future backhaul networks, their requirements, and available potential solutions. Hence, in the subsequent two sections, we discuss 5G backhaul networks and their potential technology approaches, followed by the related projects and works on joint access-backhaul.

Contributions
[C1] R. I. Rony, A. Jain, E. Lopez-Aguilera, E. Garcia-Villegas, and I. Demirkol, "Joint access-backhaul perspective on mobility management in 5G networks", IEEE Conference on Standards for Communications and Networking (CSCN) 2017, pp. 115–120, Sept. 2017. doi: 10.1109/CSCN.2017.8088608.

2.1 5G backhaul

5G is expected to employ CRAN, where most of the RAN functionalities are centralized in BBUs, and the APs, referred to as RRH, will perform only radio functionalities. Besides, traditional BS/eNB will co-exist, operating in the DRAN approach, where eNBs/BSs perform all the RAN functionalities. Additionally, as mentioned earlier, future access networks are expected to employ a large number of SCs to increase the overall spectrum efficiency and

meet the required capacity. Therefore, densely deployed cells and the density of users in the network will bring the Ultra-Dense Network (UDN) concept. To support UDN, 5G transport networks will become more complex, composed of backhaul, fronthaul, and midhaul as a consequence of the new access networks architecture. Figure 2.1 depicts the relationship among these terms.

2.1.1 Backhaul

Traditionally, the links connecting BSs/eNBs (performing RAN processing) to the core network are called backhaul links (BH), which is still a popular term. In this scenario, links connecting one BS/eNB to another BS/eNB are also considered as BH. On the other hand, in the centralized approach, i.e., CRAN, the link connecting BBU and the core network is referred to as BH. However, in 5G networks, both DRAN (RAN processing are distributed to BSs), and CRAN will co-exist and, in both cases, BH is carrying a large amount of traffic to/from the core network. For a cost-effective deployment, rather than connecting among each other, all these BSs can be connected to the core network and, thus, those BSs are connected to each other via the core network, but this adds latency. According to [11], copper wire and wireless links can be used as BH links where optical fibre has not been already deployed. Different approaches such as Resilient Packet Ring (RPR), Optical Add Drop Multiplexing (OADM) ring technology, and Wavelength Division Multiplexing (WDM) can be used for better performing BH with lower latency.

2.1.2 Fronthaul

CRAN approach centralizes most of the RAN functionalities in BBU, and the connecting links between BBU and RRH are known as fronthaul (FH). The links connecting the RRHs among each other are also considered FH links. Additionally, connecting links between eNB and the SCs acting as RRH, assuming few functionalities of the SCs are centralized into a co-located processing center with eNB, can also be considered as FH. Common Public Radio Interface (CPRI), Open Base Station Architecture (OBSAI), and Open Radio equipment Interface (ORI) are popular options for FH, although FH might have both wired and wireless links deployed. FH has already been justified as a key element of future networks, which have stringent requirements. Few novel interfaces for FH are also being explored such as fronthaul-lite, Next Generation Fronthaul Interface (NGFI), and xHaul.

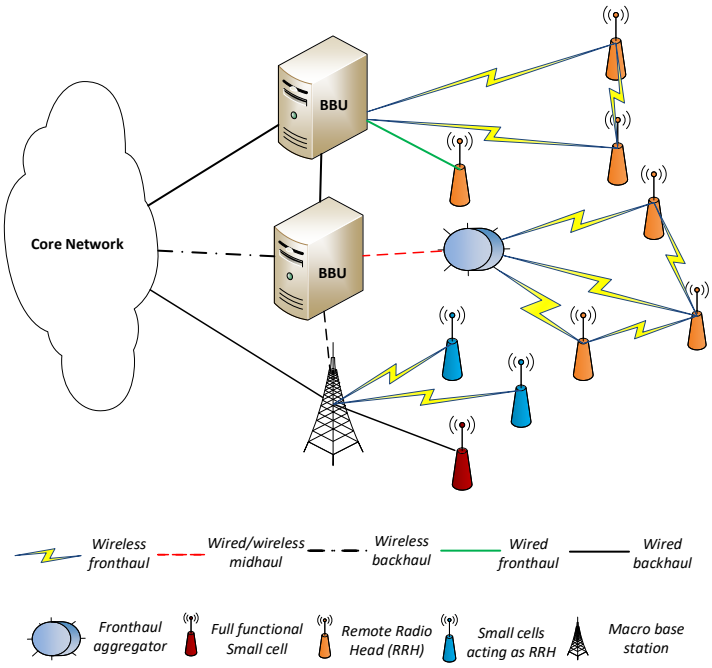

Figure 2.1: Heterogeneous backhaul in 5G networks.

2.1.3 Midhaul

In Third Generation Partnership Project (3GPP) terminology, the X2 based inter eNB interface is called the midhaul. However, with regard to future 5G based mobile networks, in [12] this term is used differently. According to [12], the link between aggregated fronthaul point and BBU is called midhaul. The idea is to achieve the multiplexing gain by hosting a few functionalities, e.g., some data compression techniques to relax the requirements for the subsequent transport network, inside the aggregator. The midhaul links can be wired or wireless links according to network requirements and availability.

Figure 2.1 depicts the heterogeneous transport network of 5G and helps to understand the separation between fronthaul, midhaul, and backhaul.

The discussed complex transport network will become a major cost component, and it will be one of the most dominant parts of the network performance in terms of capacity, latency, etc. Hence, re-thinking of transport network solutions becomes necessary. Note that hereafter the term backhaul is used as a generic term to refer to the entire transport network (including fronthaul and midhaul), although in a few cases they are also used separately when required.

2.2 Potential backhaul options

As discussed in Section 2.1, the transport network described in Figure 2.1, is one of the most challenging aspects of 5G in terms of cost, heterogeneity, and complexity. There are a few popular technologies already available for backhauling the previous generations of cellular networks. However, unlike previous generations, in 5G, the transport network includes another major element, FH, demanding a huge capacity, which will be very challenging to provide. In this section, we explore the available and upcoming technologies considered as potential candidates for future BH/FH networks.

2.2.1 Wired options

Optical fibre is, arguably, the best wireline BH solution, which can provide enormous capacity [13] and is considered as a firm candidate for future FH deployments. The basic fibre deployment is based on Passive Optical Network (PON) technology, which enables a single fibre to serve multiple ends utilizing splitters. Throughout the years, PON was evolved towards Broadband PON (B-PON), Gigabit PON (G-PON, ITU-T G.984), Ethernet PON (EPON), 1 Gigabit EPON (1G-EPON), and 10 Gigabit EPON (10G-EPON), which are based on Time Division Multiple Access (TDMA) [11]. Another approach is WDM-PON (ITU-TG.989.2), where each Optical Network Unit (ONU), located near to the end-user, exclusively utilizes a single wavelength pair to communicate with Optical Line Termination (OLT), located at the central office. WDM technology allows the utilization of a single fibre for both Uplink (UL) and Downlink (DL) communications, thus allowing the utilization of the same physical structure, while a particular wavelength is used for specific links [14]. Additionally, GPON evolved towards Next Generation PON (NGPON, also known as XG-PON), which provides a downstream data rate of 10 Gbps and an upstream of 2.5 Gbps [13].

Later, combining TDMA and WDM, TWDM-PON (ITU-T G.989) was considered, where 4 to 8 wavelength channels are allocated in each direction, and each channel is shared among a number of ONUs through TDMA. TWDM-PON was utilized and standardized by ITU-T as Next Generation PON 2 (NG-PON2), which offers 40 Gbps in downstream and 10 Gbps in upstream covering distances of up to 40 Km [13].

According to [15], Dual Fibre Coarse Wavelength Division Multiplexing (CWDM) is also a potential option for FH deployments that require a capacity of 5 Gbps. International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) has defined 18 CWDM channels with 20 nm channel spacing, where one channel is devoted to link supervision and another channel can be devoted to transporting local alarms. CWDM-based PON is simple and cost-efficient. However, it generally requires two fibres (use of CWDM allows

single fibre deployments at the cost of a reduced number of channels) and, hence, only part of the existing Fibre To The Home (FTTH) infrastructure could be directly re-used. On the other hand, Dense WDM (DWDM) solutions, which provide better spectral efficiency than CWDM, were presented in [15]. DWDM is capable of providing 10 Gbps transmission utilizing tuneable lasers [15] and enables energy efficient network design utilizing Reconfigurable Optical Add Drop Multiplexers (ROADMs) [16]. However, WDM transmitter, control, and management of the wavelengths in DWDM system are costly and complex.

In 2015, IEEE Standard Association approved and published IEEE 802.3bm-2015 [17], where 100 Gbps technology was introduced in addition to the 40 Gbps Ethernet. Aforementioned standardization amendment includes physical layer specifications and management parameters for 100 Gbps operation over multimode fibre (100GBASE-SR4), which is capable of 100 Gbps operation over the fibre. According to [17], 100 Gbps Media Independent Interface (CGMII) can be used to connect 100 Gbps capable MAC to a 100 Gbps PHY layer. Additionally, 5GPPP projects, e.g., 5G-XHaul also included 100G (100GBASE-SR4) Ethernet-based connection as a potential candidate for future fronthaul network, which is expected to provide 100 Gbps data rate utilizing base parameters of novel optical access and digital processing architectures for future mobile backhaul [18].

Copper line-based technologies, i.e., Asymmetric DSL (ADSL) and Very high-speed DSL (VDSL), have also been popular wired options for backhauling wireless networks in the past. However, they suffer from insufficient bandwidth to be considered for future mobile backhaul/fronthaul solutions. On the other hand, they offer the advantage of employing the widely deployed fixed telephone infrastructure. Additionally, G.FAST (ITU-T G.9701) can deliver up to 1 Gbps over short-distance copper lines. NOKIA prototype XG-FAST can reach up to 10 Gbps over a few tens of meters [16].

Although the literature also discusses technologies such as Space Division Multiplexing (SDM) and multi-core fibres, in this section we have focused on the most commonly referred approaches.

To summarize, optical fibre network is the best wired solution for backhauling/fronthauling modern and forthcoming mobile access networks in terms of capacity and latency. This option is more cost-effective in case the infrastructure is already available, which is not the common situation worldwide; only sixteen countries in the world have more than 15% coverage of FTTH available [1]. Furthermore, the new deployment of fibre largely increases the CAPEX. As 5G promises ubiquitous coverage, depending only on fibre-based backhaul is not an acceptable option. Thus, 5G networks have been proposed to be deployed employing heterogeneous backhaul with both wired and wireless options. According to [19], wireless technologies are serving more than 50% of the total mobile backhaul networks worldwide, and

are being considered as a very economical solution for future deployments, as well. Hence, it is also important to discuss the potential wireless options for future FH networks.

2.2.2 Wireless options

Different types of wireless options exist for fronthaul networks. These wireless technologies differ from each other basically in terms of the frequency band used, which determines channel properties such as available bandwidth and propagation characteristics. All of them have a common advantage, they all minimize the need for wires and thus, deployment is easier, faster, and cost-effective, but they have their own shortcomings, too.

Free Space Optics (FSO) uses invisible beams of light, which provide optical bandwidth connections with multi-Gbps data rates [12]. FSO uses the same transmission wavelength as fibre optics, and is usually not licensed. Moreover, FSO propagates in free space using a very narrow beam, which creates almost no interference. However, this narrow beam also requires very careful design and implementation. Performance of FSO is affected by visibility and weather conditions such as fog, snowfall, and rain, making it unsuitable for long distances or Non Line of Sight (NLoS) links.

Microwave-based wireless links normally operate in the licensed spectrum from 6 GHz to 42 GHz. This technology has been very popular for connecting rooftop BSs, and it generally requires Line of Sight (LoS). Microwave-based FH can be deployed in both Point-to-Point (PtP) and Point-to-multi-Point (PtmP) fashion while providing maximum upstream/downstream throughput in the scale of 1 Gbps [12]. As it operates on a licensed spectrum, it is an expensive option compared to unlicensed solutions. However, moderate pathlosses make it better for larger distance (tens of kilometers) communications with highly directive antennas [13].

Sub-6 GHz (e.g., licensed 3.5 GHz and unlicensed 2.4/5.8 GHz) can be used for both PtP and PtmP scenarios. This solution can be deployed as NLoS category with carrier frequencies below 6 GHz. Better propagation characteristics make Sub-6 GHz a more attractive solution in front of the microwave for FH indoor small cells with lower bandwidth requirements. However, the provided capacity will not meet the requirements of the FH, in most cases.

mmWave technology, currently operating in two different bands, i.e., E-band (70/80 GHz), and V-band (57-71 GHz), is starting to appear as a candidate for wireless FH technology since it offers a very large capacity compared to other wireless options [19]. Because of its high capacity and low-cost deployment, mmWave is becoming a very attractive option for future FH networks and is already being considered as one of the key enablers of 5G.

However, these high frequencies suffer from larger pathlosses. Additionally, their energy

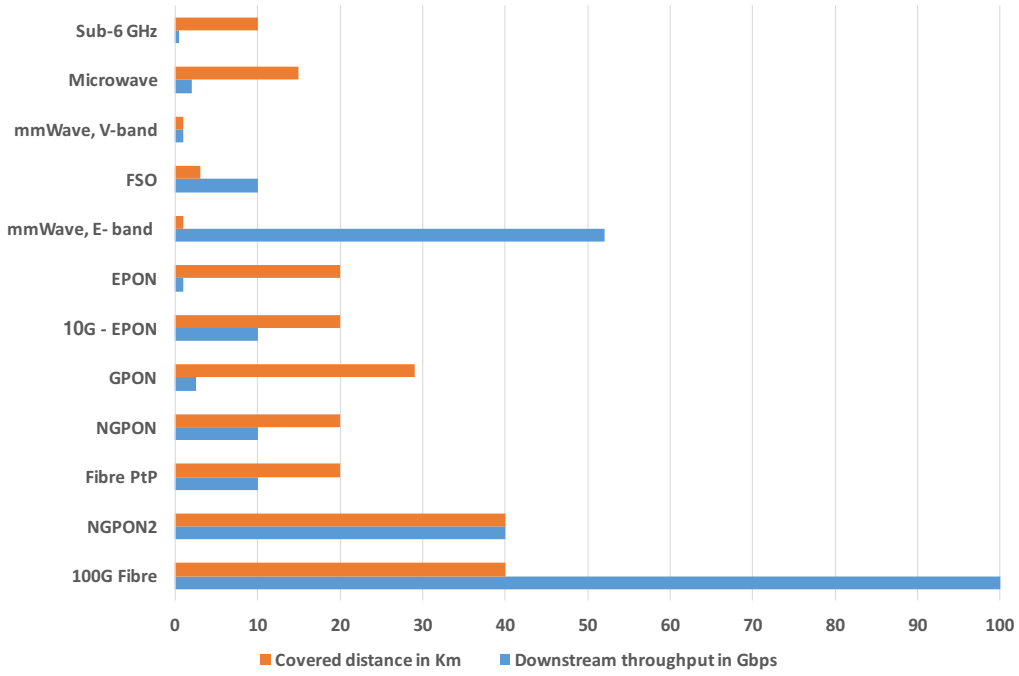

Figure 2.2: Potential options for BH/FH networks.

absorption due to atmospheric phenomena becomes prominent when the frequency range goes up to the millimeter level. On the other hand, with the use of MIMO techniques (e.g., beamforming or spatial multiplexing), those effects can be mitigated.

Reference [19] mentions two more new bands, namely W-band (92-114.25 GHz) and D-band (130-174.8 GHz) for mmWave-based communications, which are yet under development and require regulation at national/international level. E-band, which has the lowest dependency on the environmental effects among other mmWave bands, is envisioned to provide more than 50 Gbps data rate, while covering less than 1km distance in the FH network [19].

Figure 2.2 provides a comparison among different wired and wireless technologies in terms of coverage range and achievable throughput.

In future networks, a diversity of user requirements will exist leading towards eight different 5G service use cases [20] [21]. According to [21], the eight different uses cases are the following: Broadband access in dense areas (high data rate and low latency); Broadband access everywhere (50+ Mbps everywhere); Higher user mobility (500 Kmph); Massive IoT (highly reliable and secure network); Extreme real-time communications (critical reliability, ultra-low latency, robust communication); Lifeline communications (very high level of availability, ability to support traffic surges); Ultra-reliable communications (extremely low latency, zero or low mobility, high reliability, security) and Broadcast-like services (regional reach from 1 to 100 Km). Each use case presents different requirements in terms of latency,

capacity, mobility, and reliability. Hence, in 5G, APs serving different purposes will make their location more diverse and unpredictable, being not possible to feed them by wired backhaul always. Thus, wireless backhaul options become more popular and cost-effective from the operator's point of view.

2.3 Related projects and their prospects

Joint operation of backhaul and access networks is a key enabler for 5G and is considered a key research topic. Some research groups are already working on this topic and have some conclusive proposals, yet providing open issues to work on. In this section, related 7th Framework Programme for Research and Technological Development (FP7) and Horizon 2020 projects and their proposals are shortly discussed. Their contributions are further elaborated in subsequent sections.

1. **BuNGee** (1 January 2010 - 30 June 2012):

   BuNGee FP7 was a project targeting to facilitate at least 1 Gbps/km^2 throughput level in a cost-efficient way [22]. To achieve the targeted goal, BuNGee worked on a few invoking paradigms, i.e. joint design of access and backhaul mechanism, ultra-dense below rooftop radio network deployment, innovative antenna technologies, and network and distributed MIMO. BuNGee introduced aggressive joint access and backhaul scheme earlier in 2010 – 2012 [12], where a Time Division Duplexing (TDD) based frame was proposed with a mixed structure of "transmission and silent" mode and "simultaneous operation mode". In [22] BuNGee also proposed an extension of cognitive cycle (a cycle of finding available channel to be used in three steps, i.e. spectrum sensing, spectrum analysis and spectrum decision) through docition (a process of teaching, while facilitating learning) for a better performing Radio Resource Management (RRM). ETSI used the outcome of this project on its technical report on Broadband Radio Access Networks (BRAN) [23].

2. **iJOIN** (1 November 2012 - 30 April 2015):

   iJOIN is also a FP7 project. According to [12], iJOIN identified the joint operation of backhaul/RAN as a prime enabler of future mobile networks. According to iJOIN, the border between RAN and backhaul must be blurred in future mobile networks, which require the joint design of backhaul and RAN. iJOIN invented the term RAN as a Service (RANaaS), which enables a flexible RAN architecture that is neither fully centralized nor fully distributed. iJOIN successfully came up with interesting proposals

for joint operation while actually removing the border between RAN and backhaul. iJOIN also proposes the joint operation with SDN controller with user and Data Plane (DP) decoupled. From the resource management point of view, iJOIN proposed the cooperation between RAN and backhaul in two different categories, backhaul/RAN awareness and joint functional design of RAN/backhaul; later sections provide more discussion on this.

3. **MiWEBA** (1 June 2013 - 31 May 2016):

   The joint European Japanese research project Millimetre-Wave Evolution for Backhaul and Access (MiWEBA) ended in May 2016 proposes mmWave for backhaul, fronthaul, and in the access links. MiWEBA provides detailed calculations of link budget for both backhaul and access links employing mmWave [24].

4. **5G Crosshaul** (1 July 2015 - 31 December 2017):

   5G Crosshaul is one of the 5G infrastructure Public Private Partnership (5GPPP) phase 1 projects working with a focus on integrated backhaul and fronthaul transport network. 5G Crosshaul proposes unique network architecture with high-capacity switches and heterogeneous transmission links. This project identifies the requirement for 5G networks and defined their own enablers towards 5G. A network architecture with new elements and functionalities is described in [25]. 5G Crosshaul takes into consideration SDN/NFV based architecture for future mobile communications. This project already proposed a separate design of data and control plane with SDN controller in the network for integrated backhaul and fronthaul. One interesting finding from 5G Crosshaul project is that their proposed data plane architecture allows parallel operation of Packet Switch and Circuit Switch in XFE (5G Crosshaul Forwarding Element).

5. **5G NORMA** (1 July 2015 - 31 December 2017):

   5G Novel Radio Multiservice adaptive network Architecture (5G NORMA) is one of the 5GPPP phase 1 projects aimed at developing a novel mobile network architecture, which provides the necessary adaptability to handle future traffic demand. Like other future architecture proposals, 5G NORMA also adopts SDN as a network controller and proposed SDN functional blocks to control the network functions. 5G NORMA architecture enables multi-tenancy, multi-RAT, NFV, and network slicing. In [26] 5G NORMA proposes a functional architecture enabling multi-connectivity in the access network. In [27] 5G NORMA describes an algorithm for elastic resource management considering virtual cells/cell-less architecture paradigm.

6. **5G XHaul** (1 July 2015 - 31 December 2018):

   5G Xhaul is also one of the 5GPPP phase 1 projects. 5G-XHaul project worked on developing novel converged optical/wireless architecture and network management algorithms. This project also takes into account the introduction of advanced mmWave and optical transceivers and control functions [28]. 5G XHaul project concluded successfully with a demonstration of the proposed 5G Xhaul architecture, which supports the convergence of FH and BH services.

7. **5G-PICTURE** (1 June 2017 - 31 May 2020):

   5G-Programmable Infrastructure Converging disaggregated neTwork and compUte REsources (PICTURE) is a 5GPPP phase 2 project. Taking the benefit of advanced wireless and optical network solutions, 5G-PICTURE project started with a vision to develop converged fronthaul and backhaul infrastructure. In [29], 5G-PICTURE presented an idea, where an LTE-based eNB self-backhauls itself to another LTE-based eNB through an LTE connection. The presented results show improvement in network coverage and throughput. In [30], to carry fronthaul and backhaul traffic in an integrated way, an Ethernet optical channel was investigated. 5G-PICTURE also developed SDN-based Wi-Fi Small Cells with Joint Access-Backhaul and Multi-Tenant Capabilities (SWAM), an infrastructure that supports multi-tenancy, mobility, and integrated access and backhaul capabilities [31]. In [32], 5G-PICTURE proposed and evaluated an SDN-based Wireless Backhaul solution for Dense 4G/5G Small Cell Networks (SODALITE) to tackle two challenges, i) high cost and unavailability of wire-line backhauled sites, and ii) the complexity of managing wireless backhaul nodes.

8. **5G-COMPLETE** (1 November 2019 - 31 October 2022):

   5G-A unified network, Computational and stOrage resource Management framework targeting end-to-end Performance optimization for secure 5G muLti-tEchnology and multi-Tenancy Environments (COMPLETE) is an ongoing 5GPPP phase 3 project, which aims to propose a unique converged architecture to merge fronthaul, midhaul, backhaul, and 5G New Radio (NR) into an Ethernet-based common platform.

9. **TERAWAY** (1 November 2019 - 31 October 2022):

   TERAWAY is a 5GPPP phase 3 project with a vision to investigate Terahertz technology for ultra-broadband and ultra-wideband operation of backhaul and fronthaul links in systems with SDN management of network and radio resources. TERAWAY aims to go beyond mmWave and develop Terahertz-based transceivers to operate from 92 up to 322 GHz frequency band while offering data rates up to 241 Gbps.

2.4 Joint access-backhaul mechanism

From the previous discussion, it is evident that future backhaul networks will be very complex, heterogeneous, and challenging to design. Moreover, the backhaul network, connecting access to CN, will have stringent QoS requirements. Joint operation of access and backhaul, a prime enabler of 5G, comes up with alluring solutions, where backhaul and access networks work very closely to each other, thus blurring the separation line between them. In this way, backhaul and access networks become fully aware of each other's requirements and limitations. In recent research, it is commonly agreed that joint operation of access and backhaul is inevitable in future networks [33] and, subsequently, research efforts addressing different aspects of the joint operation mechanisms are available. In the following discussion, the related works are categorized as joint functional design and joint optimization of access and backhaul networks.

2.4.1 Joint access-backhaul functional design

Traditionally, in DRAN approach, BSs/eNBs perform all the RAN functionalities (e.g., admission control, resource management, MAC, etc.). Thus, the backhaul links connecting the APs to the CN are less capacity and latency demanding. However, DRAN lacks performance in terms of resource management, cost, and energy efficiency [12].

Conversely, in CRAN, BBU is responsible for the central resource management of the RRHs connected to it, enabling centralization gains. In CRAN, designing low-cost RRH is easier, and centralized control ensures efficient use of resources utilizing network-wide knowledge. This makes CRAN very cost-effective and an attractive option for operators. Additionally, according to extensive simulation results (e.g., [34]), CRAN outperforms DRAN in almost every KPI, i.e., cell-average and cell-edge throughput, spectral and energy efficiency. However, despite the benefits of CRAN, this mechanism demands very high capacity and very a low latency transport network [35].

To take advantage of both RAN approaches, i.e., CRAN and DRAN, a trade-off between these two is required, which leads to flexible functional splits. In this approach, RAN is neither DRAN nor CRAN, but flexible functional splits can be proposed, which can merge CRAN and DRAN without loosing their respective advantages. In other words, the degree of centralization of the radio access network can be flexible as a trade-off solution between CRAN and DRAN. The splitting decision can depend on the available resources. If more backhaul resources are available, a higher number of functions can be centralized, whereas this number can decrease if the backhaul network is resource-limited [36].

Such a trade-off solution, known as RANaaS, is proposed by iJOIN project [35], which

allows flexible RAN architecture. Flexible-RAN (FLEX-RAN), as shown in Figure 2.3(a), allows the RAN functionalities to transition between fully centralized and fully decentralized architectures on-demand, making RAN more flexible while also relaxing backhaul requirements. Amongst the multiple possibilities for FLEX-RAN architecture, studies [12], [37], [38] suggest PHY layer functional splits from a joint access-backhaul design point of view.

Functional splits at the PHY layer are a key enabler of the Flexible RAN concept that allows the centralization of upper layers, i.e., MAC, Radio Link Control (RLC), etc, while relaxing the stringent throughput requirements in FH network. Authors in [12] [38] [1] [39] [40] discuss functional splits at the PHY layer, and the derived trade-off between centralization and FH requirements. As depicted in Figure 2.3(b), in this work we focus on four potential splitting points at the PHY layer. As in [41], we assume that transfer of the analysis in the Uplink (UL) scenario to the Downlink (DL) scenario is straightforward and, henceforth, we only discuss the UL communication. There are four potential splitting points at the PHY layer (Figure 2.3(b)).

- In *Split-A*, all the PHY layer functionalities along with upper RAN layers are centralized and processed at BBU, resembling the CRAN approach. As mentioned earlier, in this case, all the IQ data is forwarded from the APs, also known as RRHs in a full centralization approach (Split-A), to the BBU in the UL after Analog to Digital (AD) conversion, and usually, it is referred to as Radio-over-Fibre (RoF), used in CPRI standard [40]. Thus, the FH data rate requirement for this split depends on the number of AD converter chains, sampling frequency, transport overhead, and resolution of the time domain quantizer. Hence, FH data rate requirement for Split-A is static, agnostic to the real traffic scenario, and extremely high (4.9 Gbps expected for Sub-6 GHz access technology, and 199.5 Gbps for high mmWave) [39]. A major benefit of this split is that almost no processing in AP is required, resulting in more cost-efficient APs. Additionally, even if a fully functional AP acts as RRH with Split-A, the power consumption by the APs decreases considerably since the Digital Unit (DU) in the APs can be totally powered off. Moreover, this split does not restrict any type of centralization, e.g., implementation of Coordinated Multi-Point (CoMP), enhanced Inter-Cell Interference Coordination (eICIC), which can improve the network performance [38].

  For Split-A the required FH data rate can be calculated using Eq. 2.1, where N_L presents the number of ADC chains, f_S is the sampling frequency, N_{QT} is the resolution of the time domain quantizer, and γ is the transport overhead [39].

$$C_{\text{Split-A}} = 2 * N_{\text{L}} * f_{\text{S}} * N_{\text{QT}} * \gamma \tag{2.1}$$

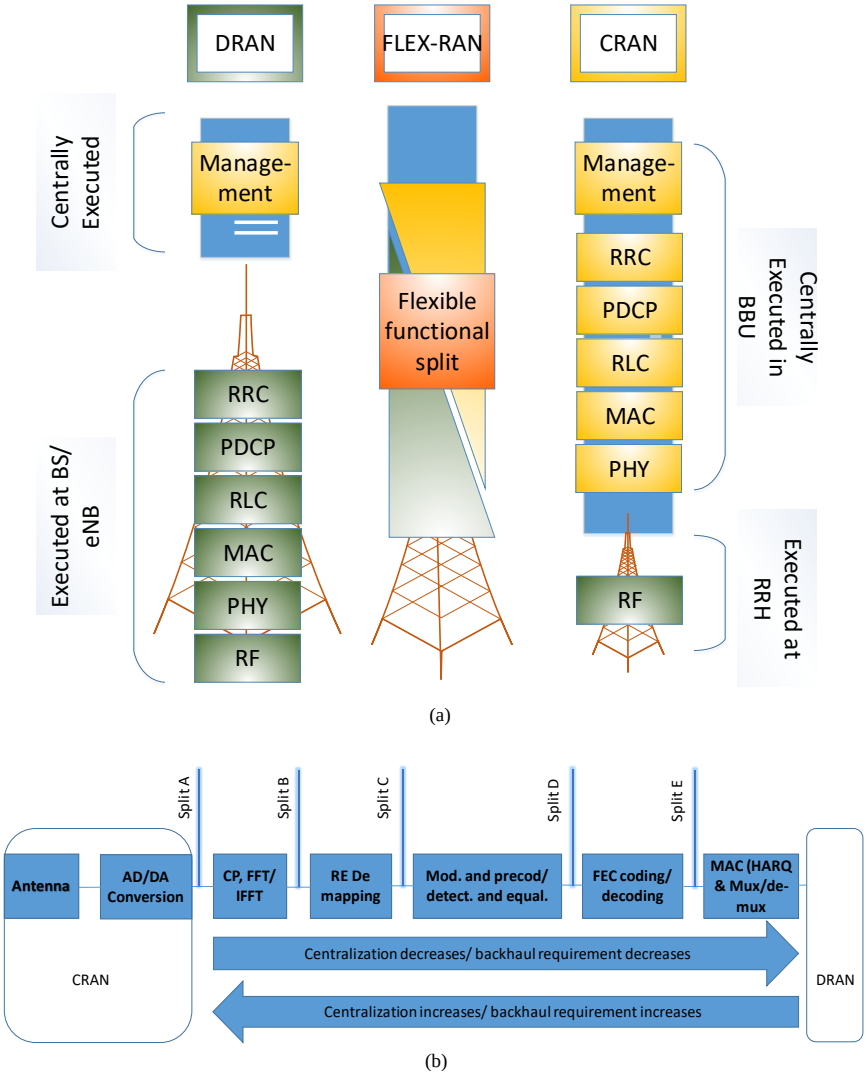

Figure 2.3: Joint access-backhaul functional design: a) FLEX-RAN, a trade-off between CRAN and DRAN; b) functional splits at PHY layer (UL/DL).

- In *Split-B*, the received signal is forwarded from the APs to the BBU in the frequency domain after Fast Fourier Transform (FFT) [40]. Thus, the required data rate is now dependent on the number of active subcarriers, the number of symbols per frame, the resolution of frequency domain quantizer, and the frame duration [39]. Despite the decreased capacity requirements, this split also has almost no restrictions on the centralization gain, but it imposes additional functionality at the APs, which increases the OPEX of the APs. Moreover, similar to Split-A, FH data rate requirement for Split-B is also static and agnostic to the real traffic.

  For Split-B, the required FH data rate can be calculated using Eq. 2.2, where N_{SC} presents the number of active subcarriers, N_{SY} is the number of symbols per frame, N_{QF} is the resolution of the frequency domain quantizer, T_F is the frame duration [39].

$$C_{\text{Split-B}} = 2 * N_{\text{L}} * N_{\text{SC}} * N_{\text{SY}} * N_{\text{QF}} * T_{\text{F}}^{-1} * \gamma \tag{2.2}$$

- *Split-C* can be considered a more practical split, since the required data rate scales with the actual data traffic in the access network. As depicted in Figure 2.3(b), the resource mapping/demapping is done locally at the APs, and hence, only the utilized Resource Elements (RE) are forwarded in the UL transmission to the BBU. Required data rate of this split is almost the same as for Split-B with additional dependency on data traffic/RE utilization factor [40]. Hence, in full load, the requirements for Split-C and Split-B are equal.

  For Split-C the required FH data rate can be calculated using Eq. 2.3, where μ is the utilization of the subcarriers.

$$C_{\text{Split-C}} = 2 * N_{\text{L}} * N_{\text{SC}} * N_{\text{SY}} * N_{\text{QF}} * T_{\text{F}}^{-1} * \gamma * \mu \tag{2.3}$$

- In *Split-D*, all the PHY layer functionalities are performed locally at the APs and the upper layer functionalities are centralized. In this split, the required data rate shows dependency on the coding rate and the number of modulation symbols [39], hence relaxing the required data rate further more. This split does not allow joint transmission and reception in CoMP [39], yet the centralization gain from joint scheduling, interference coordination and path management techniques are still achievable [38].

  For Split-D the required FH data rate can be calculated using Eq. 2.4, R_C presents the coding rate and M_{MSC} is the number modulation symbols.

$$C_{\text{Split-C}} = 2 * N_{\text{L}} * N_{\text{SC}} * N_{\text{SY}} * R_{\text{C}} * \log_2 M_{\text{MSC}} * T_{\text{F}}^{-1} * \gamma * \mu \tag{2.4}$$

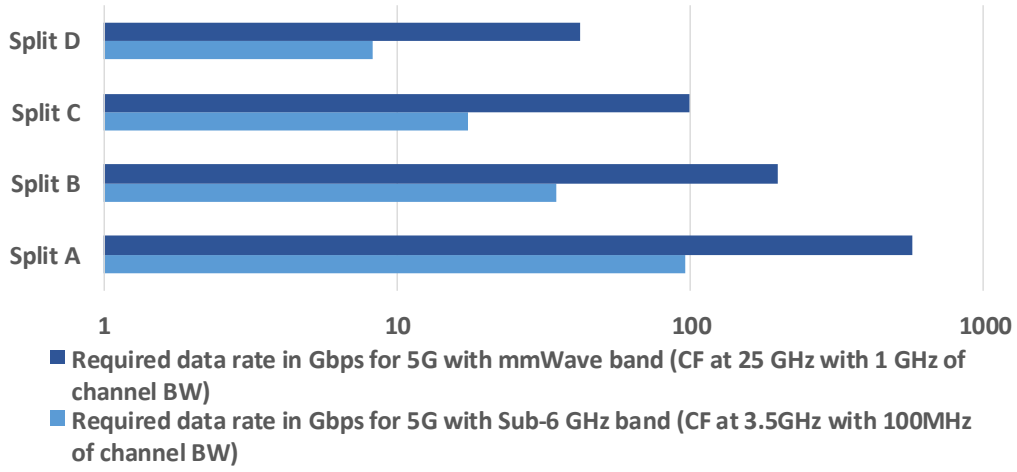

Figure 2.4: FH data rate requirements for different splits with different RATs considering 50% traffic load.

Reflected from the aforementioned discussion, different RATs in the access network have different data rate requirements in the FH. Looking towards 5G, reference [2] identifies two promising RATs, i.e. Sub-6 GHz (Carrier frequency (CF) at 3.5 GHz with 100 MHz channel bandwidth (BW)) for Macro Base Stations (MBS) and mmWave (CF at 25 GHz with 1 GHz channel BW) for Small Cells (SC). Utilizing the equations presented in [39], in Figure 2.4 we present the corresponding FH requirements for the aforementioned RATs and the different functional splits. Within the highest split point (Split-D) all the PHY layer functions are locally processed, yet upper layer (e.g. MAC and RRC) functions can be centralized to have a significant impact on the network performance [38]. Hence, the discussed functional splits at PHY layer can be potentially utilized to relax FH requirements and yet achieve a certain level of centralization benefits brought by the CRAN approach. Moreover, functional splits within the upper layers (e.g. MAC and RRC) are also possible and can further relax FH requirements, which will still pose corresponding benefits (e.g., centralized interference mitigation, enhanced spectral efficiency, load balancing) and challenges (e.g., increased computational complexity) [35].

More detail on functional split options suggested by the standardization bodies is discussed in Chapter 4.

Although, with the higher point of splits (From Split-A towards Split-D) throughput requirement for BH/FH becomes more relaxed, the latency requirement is closely tied to the user velocity [12]. In these aforementioned splits at the PHY layer, the the latency is governed by Hybrid Automatic Repeat Request (HARQ), link adaptation and scheduling

process [3]. For instance, for Split- C, latency originates from channel coherence time in DL and HARQ in UL, while for Split-B, the latency requirement is tied by the link adaptation in DL and the HARQ in UL [10]. On the other hand, latency requirement for split-D is considerably lower than for other splits, as all PHY layer functions are now locally processed at APs. However, Opportunistic HARQ, a process where the time-critical part is conducted in the APs and computationally intense part is processed at the central processing unit, is proposed by both [12] and [38] so as to relax the latency requirements over the BH/FH.

Further discussion on functional splits focused on economical cost aspects is given in chapter 4.

2.4.2 Joint access-backhaul optimization

Joint optimization of access and backhaul refers to the mechanisms, where access and backhaul networks are optimized together depending on each other's conditions. Recent research considers different approaches of such joint optimization, which are being considered as essential mechanisms for future networks.

Joint **interference management** consists in an essential technique for self-backhauling solutions, where access and backhaul networks share the same frequency band. For such scenario, in [42] authors propose two-step interference management in a full duplex system. First, self-interference (SI), which occurs due to the leakage of the transmitted signal to the receiver channel, can be attenuated into a dynamic range utilizing spatial isolation/separation between transmitter and receiver. Second, UE receiving the signal from backhaul and access nodes in the same frequency channel creates interference, which can be handled by the coordinated power control mechanism, i.e., controlling the coverage area of both access and backhaul nodes. Additionally, coordinated Multi-Point (CoMP)/network MIMO, where APs cooperate with each other to control their transmissions and receptions, can be useful to improve network performance by minimizing interference [12]. Such cooperation can also be implemented between access and backhaul nodes to perform joint interference management.

Moreover, enhanced Inter-Cell Interference Coordination (eICIC) is necessary to deal with the interference among network tiers sharing the same spectrum [43]. In eICIC, Carrier Aggregation (CA), where two or more carrier frequencies are aggregated together to support wider bandwidth, Almost Blank Subframe (ABS), a technique to stop the transmission of interferer tier in a sub-frame, and power control, are utilized to reduce interferences. Additionally, massive MIMO, beamforming, and directional antenna techniques can be further beneficial to reduce the interference in such scenarios.

Network-wide energy optimization is proposed in [44], to optimize the energy consumption together for backhaul and access networks. In this approach, switching-on/off of access or backhaul nodes considers the effect of each other utilizing an SDN controller having network wide knowledge. Additionally, reference [45], formulates the joint backhaul and access design problem by minimizing the total power consumption over the network.

Joint **load balancing** is another joint optimization approach, where load balancing is performed both for access and BH networks together, i.e., the status of the backhaul has influence over the load balancing in access networks, whereas, traditional approaches only take into account the status of the access network. Reference [46] proposes an algorithm for a TDD-based system, where tuning parameters (e.g. duration of transmission between access and BH links, achievable rate) are balanced according to the current requirements of different cells. Cell Range Extension (CRE) is commonly utilized for offloading UEs from the macro base station towards SCs in a two-tier network. In CRE mechanisms, CRE offset (CREO) is used to increase the cell range of SCs or bias the UEs to connect to the SC, as in most of the cases received power from macro BS is higher than that from SCs. However, usually, CREO does not use the information about the corresponding BH link. Reference [47] proposes a user-centric backhaul (UCB) scheme that optimizes user-cell association considering BH constraints. Additionally, [20] proposes a Multiple Attribute Decision Making (MADM) scheme for a user-cell-BH association criteria, where CREO includes three additional attributes, i.e., throughput, latency, and resilience of corresponding BH link.

RAN-BH awareness, where both the networks are aware of each others conditions, was discussed in [9], where joint routing and scheduling are implemented to select the best path according to access network requirements. Moreover, centralized connection control is proposed, where both cell load and BH capacity are considered for cell selection as RAN-BH awareness. Work in [48] is also a fair example of RAN-BH awareness, where the access network is re-designed according to the individual purpose (e.g., capacity, coverage, targeted area) of each AP, and subsequently fed with required BH resources. Reference [49] proposes a novel approach for optimizing the joint deployment of small cell base stations and wireless backhaul links. This proposal tries to find the optimal number of small cells that can be served by the constrained backhaul link in three steps. First, it checks the backhaul constraint and decides if the number of small cells can be served by the BH links. If not, small cells are disconnected starting from the one with the lowest priority. In the second step, utilizing the assumed pathloss model, it identifies the small cells with more coverage area compared to its backhaul constraint, and the coverage area for those small cells is reduced accordingly. The third step finds out the coverage holes and proposes new small cells deployment with potential available backhaul. Additionally, traffic steering and classification can be employed

by the future networks to serve the current amount of traffic, where resource allocation to the backhaul link can be a function of the current traffic.

Other works, such as [50], present backhaul aware resource allocation in the access network, while also analysing total backhaul power consumption. Backhaul aware cell association was discussed in [51], where both access and backhaul network power consumption were considered to associate UEs in an energy-efficient way. In [52], authors propose a centralized optimization technique to adjust CREO to associate BSs in a two-tier cellular network, where SCs are deployed overlaid with Macro Cells (MC)/BSs. Reference [53] balances the network load through a backhaul-aware user association technique. Additionally, [20] proposes user-centric backhaul, where CREO is also associated with backhaul network information such as latency, capacity, and resilience. In this work, authors compare the proposed user-centric backhaul with a few other approaches of user association such as Signal to Interference and Noise Ratio (SINR) based association, where SINR is the only priority for association; backhaul aware CRE, where CRE is optimized in view of the backhaul capacity; CRE with fixed bias, considering CRE bias has a fixed value, i.e., 6dB. Finally, results showed in different aspects summarize the validity of the user centric backhaul proposal.

Joint coding was proposed in [54], where available total code rate can be distributed to access and BH to decrease the overall Bit Error Rate (BER) in the network. Therefore, a strong code rate in access network and a weak code rate in the backhaul network, or vice versa, can be used, which results in a better performing backhaul, even if the backhaul channel quality is poor. In this case, data is protected even in the backhaul link by strong code employed in the access link, and vice versa.

Reference [43] discusses the necessity of a virtual spectrum pool to configure future complex networks, where many network links will employ the same frequency bands, and multiple operators may be sharing the network infrastructure. With this in mind, **joint spectrum sharing** can be another approach of joint access-backhaul optimization, where the same spectrum pool will be shared by both networks to use the scarce resources, i.e., spectrum, in an efficient way. There are many popular techniques to share the same spectrum between access and backhaul networks: i) self-backhauling, where SCs employ partitioned or entire dedicated frequency band to be used among access and backhaul links, ii) full duplex (FD) mode, where transmission and reception are performed in the same time-frequency resource, and iii) half duplex (HD) mode, where transmission and reception are performed in a separate time-frequency resource. However, each frequency re-use option introduces interference in the network that will be very challenging to optimize. Hence, we go into more detail about the available options of spectrum sharing for future networks.

In-band full-duplex (IBFD) [55] is a technique where SCs can work in full duplex mode,

backhauling themselves wirelessly with the anchor eNB and communicating simultaneously over access and backhaul using the same frequency band. In reference [55], employing FD mode, the entire available bandwidth, $2W$, is used by both eNB and SCs, and later the bandwidth is divided into W to be used for UL and W for DL communication at both eNB and SCs. Subsequently, W bandwidth dedicated for eNB UL/DL transmissions is further divided into $\eta * W$ and $(1 - \eta) * W$ for backhaul and access communications, respectively, with $0 < \eta < 1$. On the other hand, in Half Duplex (HD) mode, Frequency Division Duplexing (FDD) is used to divide the entire bandwidth among eNB and SCs as $\alpha * 2W$ and $(1 - \alpha) * 2W$, respectively, where α is usually equal to 0.5. Similar to the previous technique, allocated bandwidth W, is further divided into $W/2$ for UL and $W/2$ for DL communication (assuming α is equal to 0.5). Lastly, $W/2$ dedicated for eNB UL/DL is further divided into $\eta * W/2$ and $(1 - \eta) * W/2$ for backhaul and access communication, respectively. For both the approaches, considering that eNB serves n number of SCs, the backhaul resources are limited into $(\eta/n) * W$ and $(\eta/n) * W/2$ for FD and HD mode, respectively.

Results in [55] show that, despite the higher interference in the network, IBFD capability improves the average rates throughout the networks by a factor less than double, however, the coverage is limited to half of the obtained through FDD approach. Moreover, with the increase in SC density in the network, interference increases and, thus, worsens the interference situation for IBFD mode, whereas, the coverage probability increases for FDD mode. Additionally, self-interference cancellation technique was proposed for such interference-limited scenarios in IBFD mode, though the performance of FDD in terms of coverage probability remains better.

In reference [56], authors consider a two-tier network, i.e., macro cell/eNB tier and SCs tier, and partition the resources employing active and silent modes. In this approach, resource partitioning factor η $(0 < \eta < 1)$ is the fraction of resources on which macro cell/eNB remains silent, i.e., shuts down the transmission, and SCs serve their UEs. However, in this work, the main focus was to perform offloading from macro to SCs given that, without proactive offloading the gain from SCs deployments is very limited [56]. Hence, authors try to partition the resources by varying η to bias the offloading.

Two more approaches were mentioned in [56] for resource partitioning in such two/multi-tier networks. The straightforward approach consists in a process to search all possible UE-AP associations and allocate the time/frequency resources accordingly, which is very inefficient and computationally daunting. Another approach is the probabilistic analytical approach, where the partitioning and configuration are performed on a assumption following an certain distribution. Despite the benefit of improved average performance, in future networks, scarce resources cannot be partitioned just depending on assumptions, which might

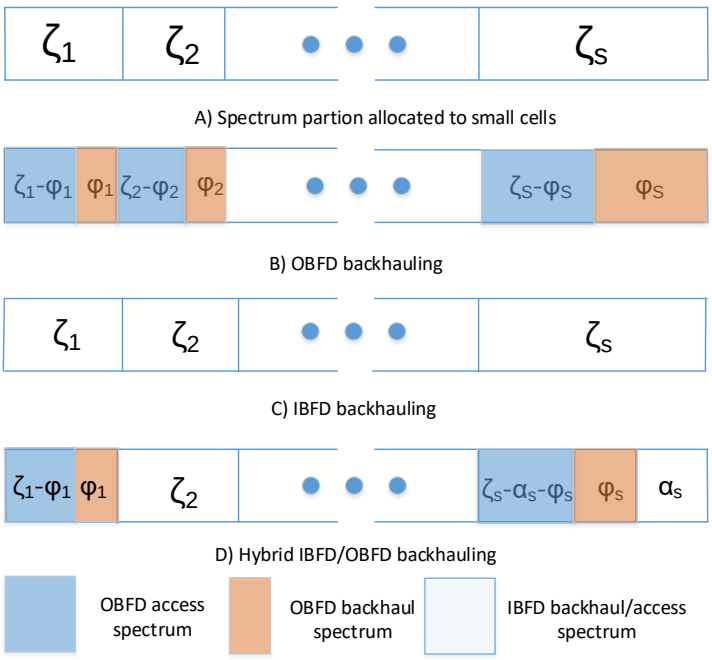

Figure 2.5: Spectrum partitioning approaches.

not be appropriate for real scenarios. Additionally, according to [57], by performing disjoint (considering access and backhaul as separate entities) spectrum partitioning, the cross-tier interference can be avoided, yet the appropriate dividing approach to be used remains an open issue. In this way, joint spectrum partitioning and allocation make the network more flexible, accurate, and resource-efficient.

In [58], authors consider a heterogeneous network, where SCs are connected to a Wireless Backhaul Hub (WBH). The available spectrum is composed of Na frequency channels for access and Nb frequency channels for backhaul links of the SCs, where $Nb = \alpha * N$, $Na = (1-\alpha) * N$, and N is the total number of available frequency channels. Accordingly, the rate of a SC, defined as the minimum among backhaul and access link rate, is mostly governed by the backhaul rate. With a small value of α, access networks having more channels cause low interference and obtain high rate availability at the access links. However, the high rate in the access links may not be supported by a small number of backhaul channels, as α is small. On the other hand, with a large value of α, access networks get less number of channels resulting into higher interference and lower rate at the access network links. Subsequently, the high proportion of the backhaul channels are wasted as a low rate in the access network does not require many channels in the backhaul links. Thus, optimal partitioning is required.

To perform the optimal partitioning of spectrum, reference [59] proposes and compares

the performance of three different approaches: Out of Band Full Duplex (OBFD), IBFD and the hybrid mode. Employing OBFD, access and backhaul transmissions use orthogonal spectrum bands, i.e. the dedicated spectrum for each SC is further partitioned orthogonally to be used by access and backhaul links. In this way, the number of channels s^{th} SC gets is $Ns = \zeta_s * N$, where ζ_s is the portion of available spectrum that s^{th} SC is assigned with and N is the total number of available channels. Ns is further divided into $\phi_s * N$ and $(\zeta_s - \phi_s)N$ and allocated for the backhaul and access transmissions, respectively, where $0 < \phi_s < \zeta_s$. On the other hand, in IBFD, the same spectrum is used by access and backhaul links in full duplex mode. In this approach, ζ_s is fully shared by access and backhaul communications, hence, the number of allocated channels for both kinds of transmissions for s^{th} SC is $Ns = \zeta_s * N$. Finally, the hybrid approach provides a flexible allocation scheme, where SC can operate completely in OBFD or in IBFD mode. In this scheme, the spectrum allocated to s^{th} SC is partitioned optimally into three portions, i.e., α_s for IBFD transmission, ϕ_s for OBFD backhaul transmission and $\zeta_s - \alpha_s - \phi_s$ for OBFD access transmission. Figure 2.5 depicts the exposed processes of spectrum partitioning schemes.

To summarize, the joint operation of backhaul and access improves the network performance to a great extent enabling a new flexible architecture with enormous benefits. This paradigm increases overall resource utilization efficiency and limits the CAPEX of the network. Functional split in different layers increases the flexibility of the network, while network requirements and available resources can be adaptive to each other's conditions. Moreover, various splitting approaches with different requirements will lead to heterogeneous backhaul solutions with different technologies adopted. Joint optimization of access and backhaul allows to share and use precious resources efficiently. RAN/BH awareness enables adapting to each other's requirements, while joint optimization can improve the overall network performance by optimizing the resources. Network-wide interference management and energy optimization take into account both access and backhaul networks together, thus improving the network-wide performance. Different techniques for the cooperative design of access and backhaul are available and can act as key performers in future networks. In UDN, the number of connected devices and of cells serving those devices will be very large. This fact leads to more complex computation for jointly optimizing RAN and backhaul networks. Nonetheless, joint access-backhaul mechanism can be a good option to handle this complexity.

Chapter 3

Access-aware backhaul optimization

As discussed in Section 2.1, designing the complex transport network in 5G will be challenging as it will require high capacity, low latency, flexibility, reliability, etc. To meet all the requirements, 5G will employ a heterogeneous transport network, where wireless backhaul options are expected to play an important role. In such a scenario, in order to meet the high capacity requirements, it is very important to utilize the available backhaul resources in a very efficient way, and thus, make the network cost-effective and resource-efficient. Although a few enablers such as FLEX-RAN and different functional splits are already identified, which can reduce the capacity requirements imposed on the backhaul, they still compromise the centralization gain to some extent. Thus, in this section, in our proposed access-aware backhaul optimization, we aim to ensure the best use of available resources without compromising the access network performance. In such an approach, access level requirements are calculated first, and required backhaul capacity is assigned accordingly.

Access-backhaul mutual awareness is one of the pillars for joint access-backhaul operation, where both the networks are aware of each other's conditions. To ensure the best use of scarce resources, this awareness is a prime requirement. The aforementioned awareness can be implemented according to many aspects as discussed in Section 2.4.2. However, most of the related works of such access-backhaul awareness optimize the access network according to the backhaul conditions or limitations. Contrary, in our proposal, we seek to optimize backhaul network capacity considering access network requirements. Traditional approaches usually offer maximum available backhaul capacity to backhaul links regardless of access networks' current requirements. Such an approach might result in an overprovisioned backhaul link and, at the same time, it might create congestion in other backhaul links due to lack of capacity. In 5G, with the higher mobility and device density, the access network

will be rapidly changing, where different APs will serve a varying numbers of users, thus, requiring dynamic backhaul capacity. On the other hand, a novel backhaul architecture is expected to offer the flexibility to accommodate those variations. Hence, in our proposal, we use this flexibility to optimize the available backhaul capacity among different backhaul links under a common controller according to corresponding access network requirements. Such an approach can be very beneficial to 5G networks as this scheme is proven to allow the efficient utilization of scarce resources.

Contributions

[C2] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Access-aware backhaul optimization in 5G." In Proceedings of the 16th ACM International Symposium on Mobility Management and Wireless Access, pp. 124-127. 2018. doi: 10.1145/3265863.3265881.

3.1 Proposed solution and analysis framework

In current deployments, all BH links are commonly offered equal and highest available capacity, which is agnostic to the current access network's requirements, making use of all available resources. In such an approach, some links may be overprovisioned and, thus, wasting valuable resources while, at the same time, it might create congestion due to insufficient resource allocation in other parts of the network. Moreover, in future dense networks, higher user density and mobility will create very unpredictable scenarios, where different APs will serve a varying number of users and, consequently, each processing unit (i.e., BBU in CRAN) might serve the different amounts of APs and the users attached to them. Therefore, BH links are expected to carry a different amount of data and, thus, require different link capacities, provided in the form of resources, which sometimes are shared with the RAN (e.g., frequency channels, time slots, etc. in the case of wireless-based BH). The aforementioned idea of access-backhaul awareness is validated in the following discussion towards an access-aware BH capacity allocation scheme for different BH links under a common central controller.

In the envisioned system, capacity allocation of different BH links depends on the current requirements of the respective served RAN. That is, access level requirements are calculated first, and corresponding BH link capacity is allocated accordingly. Let us consider an urban area of 1 km^2 in what we think represents a realistic future dense network, where eNBs co-exist with CRANs. We assume the scenario illustrated in Figure 3.1, where three BH links (black solid lines) serve three heterogeneous RAN with different capacity demands.

According to the International Mobile Telecommunications for 2020 (IMT-2020) [60], in 5G, support for connection density up to 10^6 is expected, and hence, we consider 10^6 active devices within our considered area. Among those devices, we assume[1] 100,000 are data-rich mobile devices expecting data rates between 0 to 10Gbps, whereas other 900,000 are IoT devices expecting data rates between 0 to 250Kbps [61]. We also consider different user density for each BBU/eNB: one BBU (BBU1) covers an area with larger device density (serving 30% to 50% of device population), a second BBU (BBU2) with medium density (20% to 40%), and an eNB serving the rest of the devices. This assumption ensures that corresponding BH links have different requirements. Finally, in order to be consistent in the capacity-limited BH premise discussed in the previous sections, we restrict the system with 80% BH capacity availability, that is, the system can serve, on average, 80% of maximum possible offered traffic.

For the considered system, in current deployments, the total available BH capacity provided is static, being equally distributed among the different links. On the contrary, in our proposed scheme, Control Plane (CP) and data plane are decoupled and all the BBUs/eNBs are connected to the central controller via the control plane, whereas data planes are aggregated into a BH aggregator, whereby BH resources are provisioned dynamically according to the varying demands of the different links (Figure 3.1). The central controller, which is aware of the traffic per RAN, is capable of distributing BH capacity accordingly, through the aggregation point. However, when congestion arrives at the BH, i.e., total BH capacity is not sufficient anymore, the available BH capacity is distributed in a proportional fairness basis.

For such a dense network, in a traditional approach, the backhaul links connecting the eNBs and BBUs to the core network are equal and/or have fixed capacity. On the other hand, in our proposed scheme, the data plane of the backhaul links are aggregated together into a backhaul aggregator point, whereas, the control plane is connected to the central controller. Available system backhaul capacity is shared among the backhaul links from the aggregator, which is also controlled by the central controller. The central controller, having the network-wide knowledge, also predicts the future expected load utilizing learning techniques (a technique to perform such prediction is further studied in Section 5.2). In the access-aware backhaul capacity scheme, the central controller calculates the access network requirements and distributes the available backhaul capacity among different backhaul links from the backhaul aggregator according to their corresponding requirements. In this way, a backhaul link gets the capacity it requires until a congestion point, where backhaul capacity

[1] As a reference, Manila has the highest density of population, 41,514 habitants/km^2; each of them having 2.4 devices on average would make 100,000 devices.

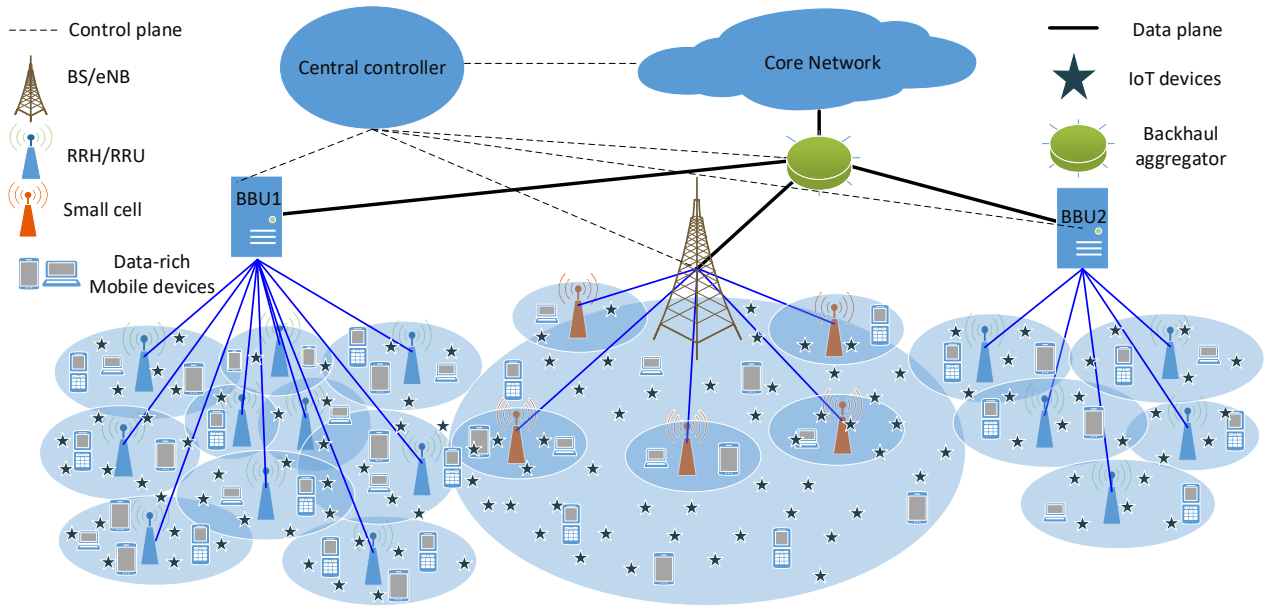

Figure 3.1: Heterogeneous 5G scenario with 2 BBUs and 1 eNB covering $1km^2$ area.

is not sufficient anymore, after which, resources are fairly distributed.

3.2 Discussion of the results

In this subsection, utilizing a few initial results, we want to demonstrate the importance of the joint access-backhaul mechanism through the presented access-aware BH optimization proposal. In Chapter 5, we present more advanced techniques for such joint resource (i.e., joint spectrum sharing, load-based spectrum allocation using machine learning techniques) mechanisms between access and BH networks.

To present the comparative results, we translated the depicted scenario (Figure 3.1) into a Matlab simulator to perform the simulation of static deployments and the access-aware BH approach.

Figure 3.2 depicts the simulation results of the presented scenario during 24 hours, in terms of Backhaul Capacity Utilization Factor (BCUF), which is the ratio of **required** versus **provisioned** BH capacity for a particular BH link. Thus, BCUF below 1 denotes that the BH capacity is not fully utilized, i.e. it is overprovisioned, whereas BCUF above 1 denotes BH link is congested, i.e. it would require more capacity. A large number of random simulations were generated, where the pattern of traffic demands and simultaneously active users follow that of real traces. Said network traces were collected in a current LTE deployment in a

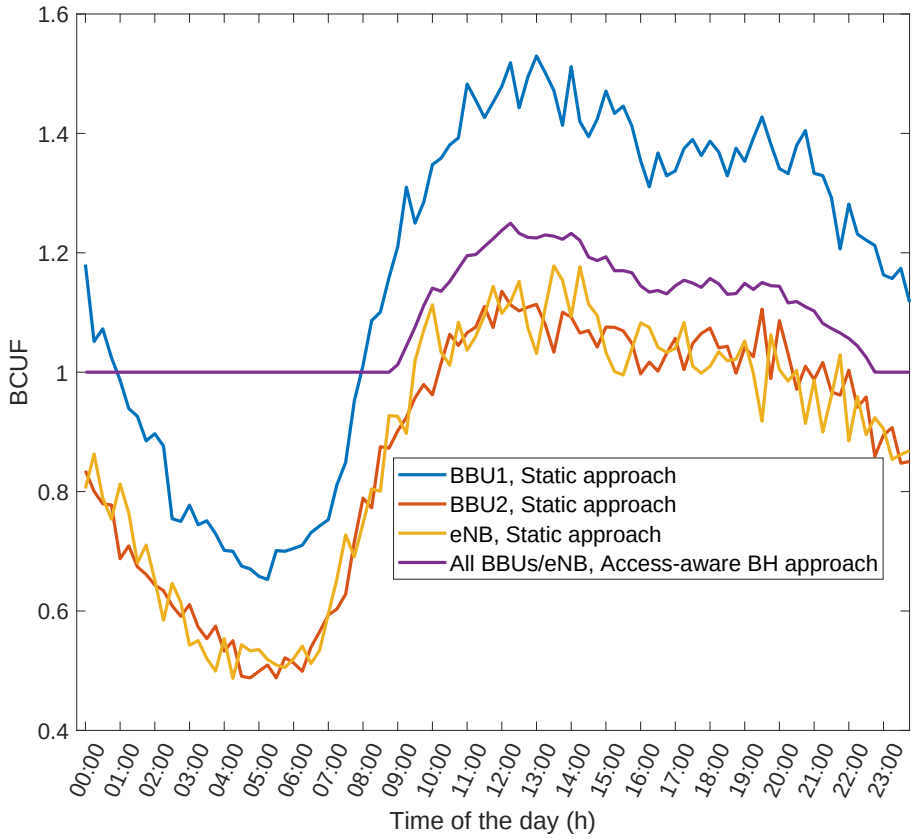

Figure 3.2: System BCUF during 24 hours with 80% system backhaul capacity.

European city over a period of two weeks; measured parameters (e.g. number of active UEs) have been scaled up to match the future 5G scenario described previously, following [60]. Results are averaged over a 24h period to ease its visualization. From Figure 3.2, in a static approach, all the BH links are overprovisioned during off-peak hours (i.e. from 02:00h to 08:15h). After that, the BH link corresponding to BBU1 experiences congestion, since it belongs to the RAN with the largest user density, while BH links corresponding to BBU2 and eNB reach the congestion point later on (i.e., around 09:45h). On the other hand, in the proposed access-aware BH scheme, provided that the load does not exceed the BH capacity, BCUF is 1 for each BH link (overlapping purple line), since BH capacity is distributed according to the current requirements, resulting in an efficient utilization until the congestion point arrives. Since we restrict the BH capacity to 80% of the maximum offered traffic, we observe a time period (i.e., peak hour), while BCUF value is more than 1, i.e., the system is congested. Evidently, during this congestion period also, the purple line is overlapping for all the three considered BH links, therefore, the congestion is also evenly distributed. Note that here we assume that the network controller has a perfect knowledge of the actual load

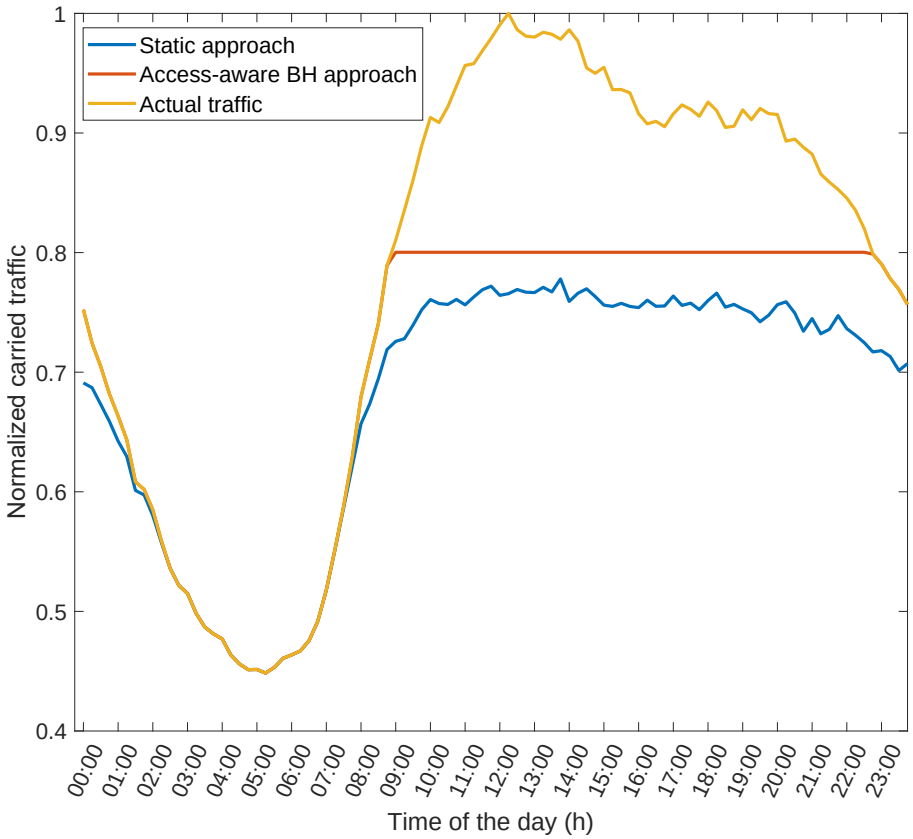

Figure 3.3: Normalized carried traffic for different approaches during 24 hours with 80% system backhaul capacity.

at each AP; in practice, a central controller would act based on predictions obtained from the constant monitoring of the network (e.g., [62], Section 5.2). In this approach, the unused or saved capacity by those links carrying less load can be distributed to other BH or access links by the central controller, if required.

Figure 3.3 depicts the normalized carried traffic (the ratio between carried traffic and maximum offered traffic) for both approaches. Evidently, the access-aware BH approach is able to carry more traffic than the static approach considering the same BH capacity condition, i.e., 80% in this case. While backhaul capacity is enough during the off-peak hour for both the approaches, the access-aware BH approach shows better performance in terms of carried traffic during the congestion period, i.e., peak hours.

Figure 3.4 represents the gain of access-aware BH approach over the static approach (in %) during 24 hours for different BH capacity conditions. When BH capacity condition is high (e.g., 70% to 90%), access-aware BH uses the whole available capacity and distributes it intelligently (i.e., according to the current requirement of each link), while the static

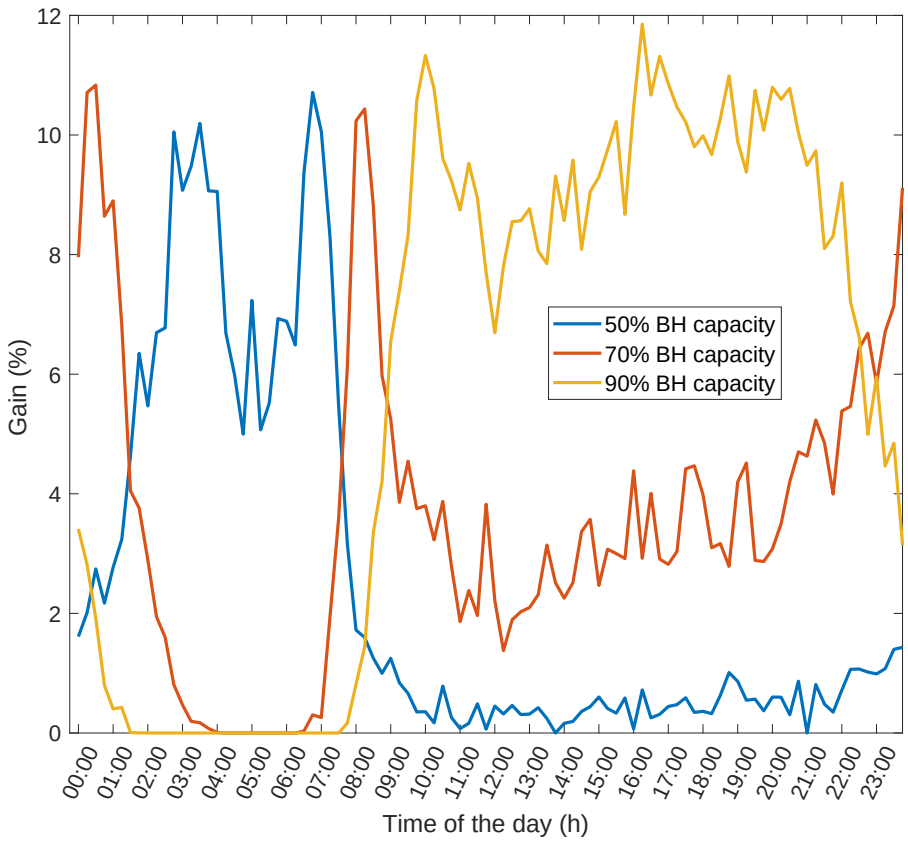

Figure 3.4: Gain of access-aware BH approach during 24 hours for different backhaul capacity conditions.

approach may have links wasting capacity. During off-peak hours both approaches provide enough capacity to serve the required traffic, and thus, no gain is observed. On the other hand, during the peak hours, gain of access-aware BH approach increases, becoming more significant the higher the BH (i.e., 90%). Conversely, when BH capacity is more restricted (e.g., 50%), the maximum gain is observed during off-peak hours. Both the approaches are out of capacity during peak hours, hence showing the similar performance (i.e., 0% gain). Also note that, when BH resources are really scarce, even during off-peak hours, some links may reach the saturation point following the static approach, while the access-aware BH approach has the means to overcome this circumstance, thus showing high gain performance. Therefore, the access-aware BH approach gets the most when the system works close to the saturation point. In case the system is extremely overloaded or really underutilized, both approaches provide similar performance.

Figure 3.5 depicts the percentage of time the system shows congestion within 24 hours for different availability of BH capacity conditions. It is clear from the figure that the access-

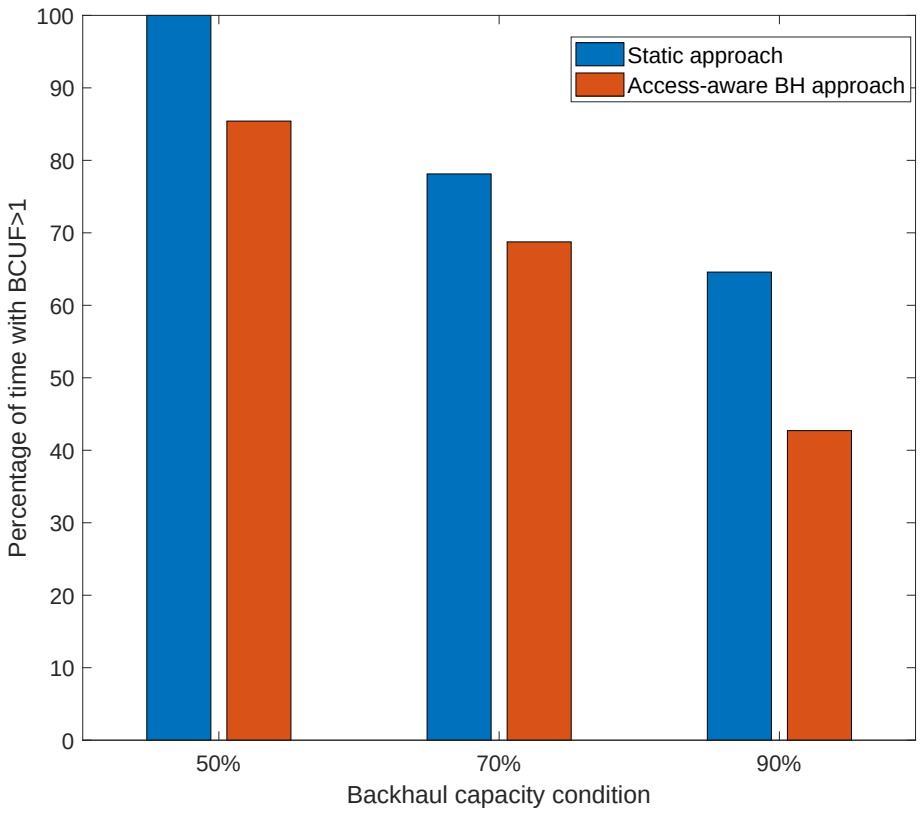

Figure 3.5: Percentage of time with BCUF>1 for different backhaul capacity conditions.

aware BH approach experiences less traffic loss (10% - 20% less time under congestion) than that with the static approach.

We also observed that the benefits of the dynamic access-aware BH approach increased as the traffic supported by the different BH links becomes more unbalanced. Obviously, when the users are evenly distributed over the different areas, a static approach based on equal distribution of resources shows the same performance as the dynamic approach.

3.3 Summary and future work

The presence of a capacity-limited BH seems a realistic assumption, which brings new challenges and requires the best usage of scarce resources. Joint access-backhaul optimization validates the dependency between both networks and facilitates the efficient utilization of resources. Presented preliminary results support the statement, in which access-aware BH optimization technique brings benefits over the static approach in terms of resource efficiency.

In this proposal, only backhaul capacity has been considered, whereas, the same mecha-

nism can be implemented in fronthaul network as well. Local controllers can be employed at BBUs and eNBs to control corresponding fronthaul networks. In this way, local controllers can calculate and learn fronthaul link requirements for each AP and share resources, i.e., total capacity can be distributed among fronthaul links accordingly. Moreover, the central controller can utilize the knowledge of local controllers to optimize the backhaul capacity distributions accordingly. In this way, employing a common controller, every backhaul link is sharing the same resource pool, and access aware backhaul scheme optimizes the distribution accordingly to ensure the best use of precious resources.

In this chapter, with preliminary results, we validate the necessity of joint access-backhaul mechanisms through the evaluation of a simple access-BH awareness method. In the later chapters, we present more developed techniques to perform joint access-BH mechanisms.

Chapter 4

Cost optimization of 5G fronthaul based on functional splits at PHY layer

5G is coming with a promise to provide ubiquitous coverage with high data rate availability. To do so, densification of access points to enhance the system capacity is anticipated. For managing such a densely populated network, 5G is expected to employ CRAN, where most of the RAN functionalities are centralized in a central processing unit. This centralization reduces operational costs and eases implementation of advanced technologies, such as, CoMP and eICIC, in a cost-efficient way.

To reach the expected network capacity, 5G will densely deploy SC, and thus, in addition to the aforementioned requirements, 5G transport network will be very complex from the architectural point of view. In CRAN, most of the functionalities are centralized in BBU, whereas APs, also known as RRHs in this architecture, are left with basic Radio Frequency functionalities [63] [64]. In this scenario, 5G transport network presents new types of links, i.e., fronthaul links connecting APs to the BBU, and Midhaul links connecting a few aggregated FH links to the BBU [12]. On the other hand, in [65], ITU proposed to split further in BBU, separating it in Distributed Unit (DU) and Centralized Unit (CU) for 5G. At this advanced stage of 5G research, such architecture composed of CU, DU, and RRH can be found in many references [66] [67]. In this scenario, the links connecting RRHs to the DU (deployed in a more distributed fashion, closer to the RRHs and performing lower layer functionalities, sometimes also acting as FH aggregator) are referred to as FH, whereas the links connecting DU to the CU, which can also act as Midhaul aggregator, are referred to as Midhaul. CUs are connected to the Next Generation Core (NGC) through the BH. In a few cases, CU and DU can be co-located, acting as CRAN BBU, hence, no midhaul is needed in the transport network (e.g., see Figure 4.1).

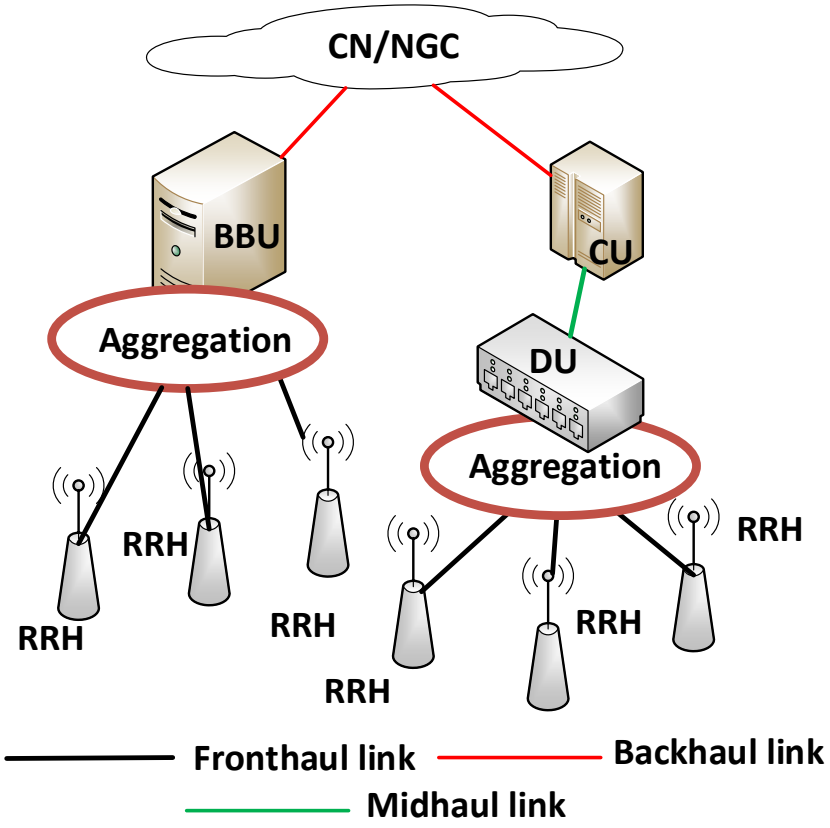

Figure 4.1: Heterogeneous transport network in 5G.

However, CRAN imposes stringent requirements on the fronthaul, i.e., the link connecting access points to the central unit, in terms of capacity and latency. Furthermore, future fronthaul networks are expected to rely on wireless technologies, since wired options are costly, not scalable and not always suitable for all scenarios [1]. Therefore, meeting the expected requirements of fronthaul network utilizing capacity-limited wireless technologies may become an inescapable bottleneck. In this chapter, we revisit different functional splits at the PHY layer in terms of data rate requirements and respective cost and discuss the combination of different splits aimed at minimizing OPEX and the overall cost, and maximizing the centralization gains, while keeping the capacity requirements below the limit of the fronthaul.

Contributions

[J1] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Cost Analysis of 5G Fronthaul Networks Through Functional Splits at the PHY Layer in a Capacity and Cost Limited Scenario." IEEE Access 9 (2021): 8733-8750. (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)). doi: 10.1109/ACCESS.2021.3049636.

[C4] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Optimization of 5G fronthaul based on functional splitting at PHY layer." In 2018 IEEE Global Communications Conference (GLOBECOM), pp. 1-7. IEEE, 2018. doi: 10.1109/GLOCOM.2018.8647928.

4.1 Cost distribution of different splits at PHY layer

In this section, we revisit the functional splits at the PHY layer, which have been studied for a long as a trade-off solution between CRAN's centralization gain and stringent capacity and latency requirements on FH links. Flexible-RAN (3GPP-TR 38.801 Rel.14) or RANaaS [68] allows shifting RAN functionalities between centralized and distributed RAN architecture. Similarly, FluidRAN [69], another form of flexible RAN considers three possible splitting points: i) Packet Data Convergence Protocol (PDCP)-RLC split, where up to PDCP functionalities are performed at CU and the rest are performed at AP, ii) MAC-PHY split, where only PHY layer functionalities are left at the AP and all the upper layer functionalities are centralized at CU, and iii) PHY split, which is equivalent to a total CRAN approach. In TR 38.801 Rel.14., 3GPPP defines different functional splits between central and distributed unit. In [65], ITU mapped the functionalities between CU and DU, following the functional splits recommended by 3GPP.

Table 4.1: Cost analysis of AP and BBU, the dependency on PHY layer splits [1].

Splits	**Benefits**	**Drawbacks**
Split-A	• Allows full centralization of functions • Maximum centralization gain • Almost no processing at AP • Cost efficient AP • Energy efficient network [12]	• Extremely high capacity requirement at FH link • Strict latency requirements • Agnostic to the actual data traffic
Split-B	• Reduced FH capacity requirement in comparison to Split-A • Almost no restriction on centralization gain [38] • Energy efficient	• Increased cost of AP in comparison to Split-A • High capacity requirement at FH link • Agnostic to the actual data traffic [C4]
Split-C	• Unutilized REs are not forwarded • Capacity requirement scales with the actual data traffic • On the basis of occupied PHY resources, statistical multiplexing gain is achievable [38]	• More functionalities at AP, hence, expensive AP • In full load, capacity requirement is the same as Split-B [C4] • Additional latency generation [38]
Split-D	• Very low capacity requirement on FH link • Less stringent latency constraints • All the PHY layer processing is performed at AP	• Less centralization gain • Joint transmission and reception in CoMP is no more possible [39] • Increased cost of AP

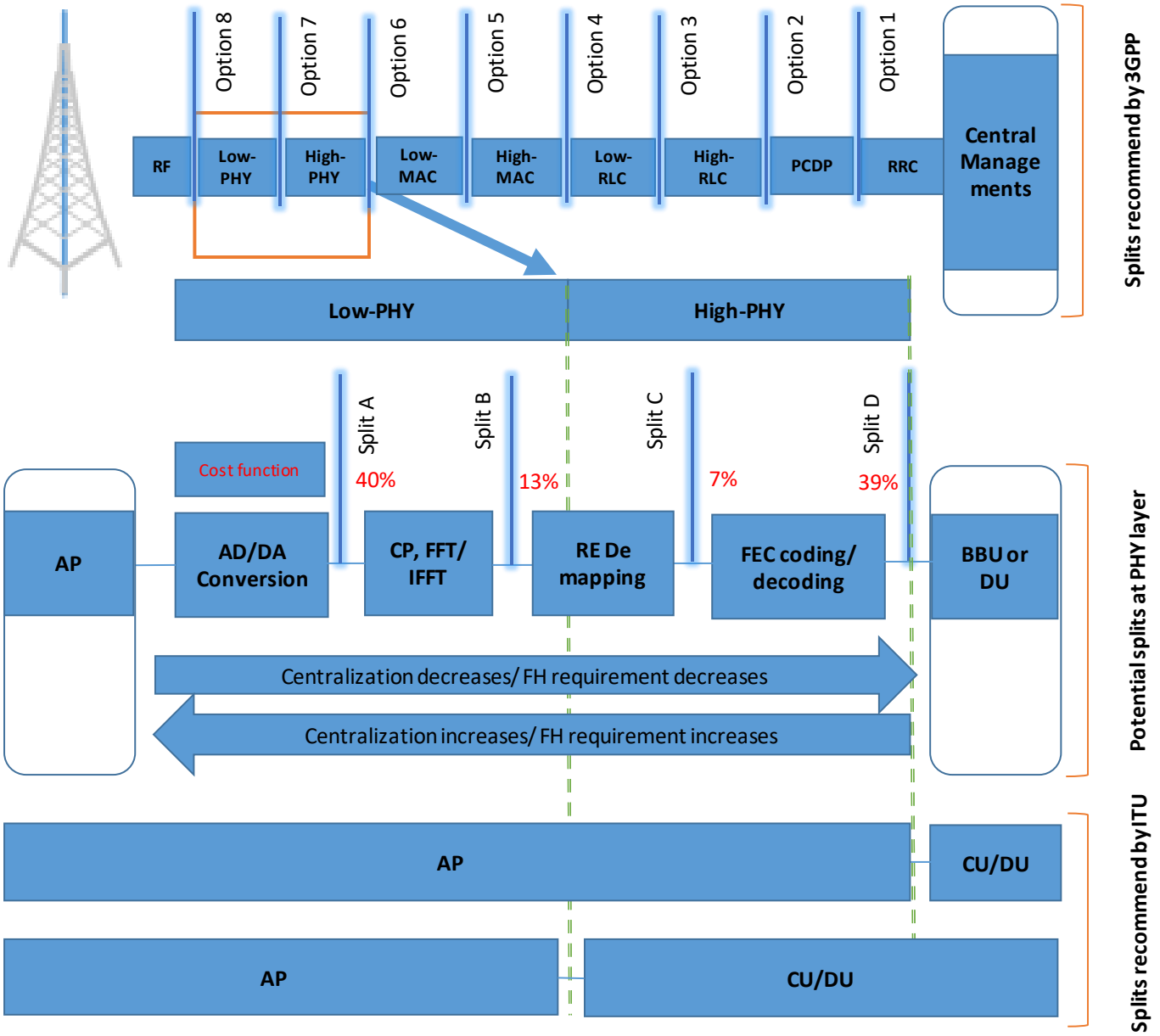

Figure 4.2: Potential splits at PHY layer and cost distribution.

On the other hand, using functional splits at PHY layer consists in the idea of finding potential functional splits within the PHY layer, where PHY layer functionalities can transition between central and distributed unit; the upper layer (e.g., MAC, RRC) functionalities are still centralized. Figure 4.2 depicts the potential functional splits at the PHY layer. Functional splits recommended by 3GPP are shown in the upper portion of Figure 4.2, where Option 7 (High-PHY) and Option 8 (Low-PHY) are the splits related to the PHY layer.

Similar splits were presented by ITU, mapped/referred to as "5G(b): low layer split" in [65] (lower portion of Figure 4.2). Concentrating on functional splits at the PHY layer, further splitting between Option 7 and 8 is presented in the middle part of Figure 4.2, where four potential options are shown. All the splits (A, B, C and D) have their own benefits and short-comings, cf. Table 4.1.

Split-A represents total centralization of functionalities or CRAN, i.e., 3GPP Option 8. Using this split, maximum centralization gain is achievable with the cost of huge capacity and latency requirements on the FH links. In Split-B, FFT is performed at the AP, which decreases the capacity requirement on the FH links [70], since the required capacity depends on the number of active subcarriers. Note that, both Split-A and Split-B are agnostic to the actual data traffic. However, if the RE mapping/de-mapping is performed at the AP, only the utilized REs have to be forwarded, hence, the capacity requirement of FH links becomes traffic dependent, and thus, Split-C can be considered as a more realistic split [C4]. Finally, Split-D decentralizes all the PHY layer functionalities, which are performed at the APs. This option represents 3GPP Option 7, and allows the centralization of all layers above the PHY. Split-D imposes the lowest capacity requirement, which is around 10% of Split-A requirement [C4].

Table 4.2: Cost analysis of AP and BBU, the dependency on PHY layer splits [1].

Parameters	**CAPEX** (€)	**OPEX** (€/year)
Macro Base Station (MBS)	MBS_{CAPEX}: 53,110	MBS_{OPEX}: 19,775
Small Cell (SC)	SC_{CAPEX}: 7,910	SC_{OPEX}: 1,950
MBS with functional splits	$MBS_{CAPEX} \times f(S)$	$MBS_{OPEX} \times f(S)$
SC with functional splits	$SC_{CAPEX} \times f(S)$	$SC_{OPEX} \times f(S)$
BBU while centralizing a MBS	$MBS_{CAPEX} \times (1 - f(S))$	$MBS_{OPEX} \times (1 - f(S))/16.34$ [71]
BBU while centralizing a SC	$SC_{CAPEX} \times (1 - f(S))$	$SC_{OPEX} \times (1 - f(S))/16.34$ [71]

4.2 OPEX-based optimization of fronthaul

As concluded from the previous discussion, utilization of PHY layer splits can relax the stringent requirements of FH network in future RAN. Nonetheless, with higher splits, centralization decreases, which affects the network performance. Additionally, according to the cost function distribution of different splits presented in [1], with higher splits, the total

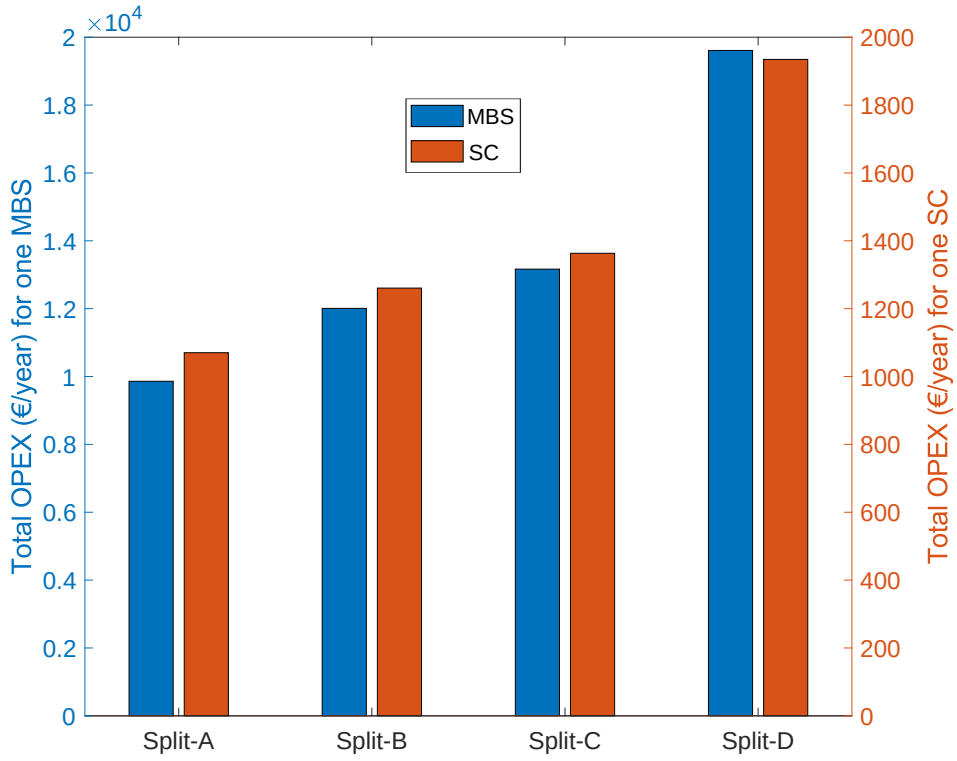

Figure 4.3: OPEX (€/year) for different split options, and for one MBS and one SC.

OPEX (i.e., OPEX of AP + OPEX of BBU) increases. Authors in the aforemetioned work distributed the cost of PHY layer functionalities among different splits as depicted in Figure 4.2[1]. In this way, OPEX of APs for different splits can be calculated as (AP_{OPEX} x *f(S)*), where S refers to the split (A,B,C or D), and AP_{OPEX} is the OPEX corresponding to an individual AP. W_s is the cost function shown in Figure 4.2 (in %), so that $f(S) = \sum^{S} W_s$. In other words, *f(S)* represents the summation of cost functions (W_s) of the functionalities run at the AP, e.g. for Split-C, *f(C)* = 0.4+0.13+0.07. Similarly, OPEX of BBU also varies for different splits as (BBU_{OPEX} x (*1* - *f(S)*).

Table 4.2 summarizes the OPEX calculation details of BBUs and APs for different splits, as suggested in [1]. Utilizing the presented values, yearly total OPEX(€) (OPEX of AP + OPEX of BBU) for one MBS and one SC are illustrated in Figure 4.3. Evidently, utilization of higher splits (i.e., Split-D/C) increases the OPEX. Thus, it is both cost and centralization efficient to utilize lower splits (i.e., Split-A/B) at the PHY layer. Additionally, from Figure 4.3, it is shown that it is more cost-efficient to operate MBS with lower splits, since the

[1]According to [1], remaining 1% of the cost belongs to the MAC layer, and it is always included in the BBU's OPEX.

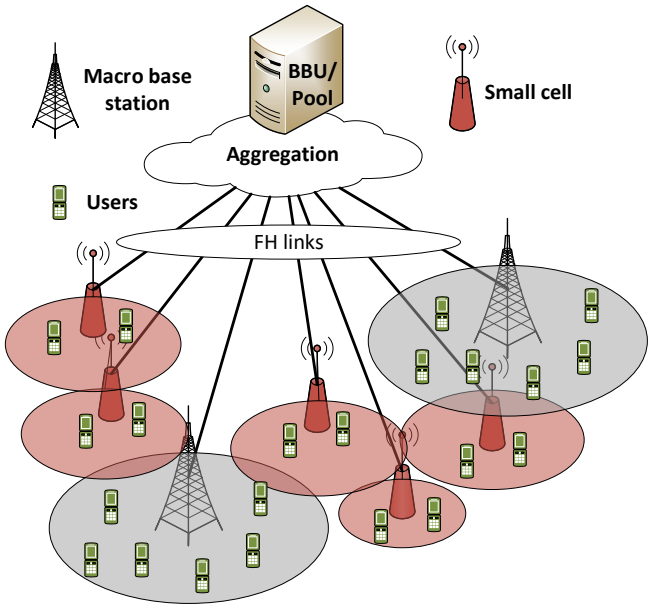

Figure 4.4: Considered scenario for evaluation.

difference of OPEX between MBS and SC is of one order of magnitude. On the other hand, as illustrated in Figure 2.4, recall that lower splits ask for higher data rate, which is a big challenge in future wireless-based FH networks.

From the above discussion, we can conclude that it is necessary to quantify the trade-off of the different splits at the PHY layer in order to find the optimal combination of splits in a deployment, which minimizes the cost, and maximizes the centralization, while keeping the capacity requirements within the limits.

4.2.1 Proposed evaluation framework

Flexible RAN, which offers a trade-off between CRAN and DRAN is a stand-up concept making the transport network more practical and flexible for future mobile networks. In this work, we focus on the PHY layer splits, considering that upper layers are centralized and processed at the BBU. In this section, we evaluate different strategies based on the concept of Flexible RAN, showing that the discussed splits at the PHY layer truly open an opportunity to make the best use of scarce resources in an efficient way to ensure the maximum achievable centralization while minimizing costs.

To perform the evaluation, we consider a dense area of 1 km^2 served by 25 MBSs[2](100%

[2]Inter Site Distance (ISD) is set to 200m following 5G-PPP's envisioned scenarios and use cases defined in [2].

Table 4.3: Parameters used for evaluation

Parameters	**Value**
Number of MBSs	25
Number of SCs	200
Traffic load in MBS	100%
Traffic load in SC	50%
Available capacity in BBU/FH aggregator	10% - 90% of the aggregated requirement for 225 all Split-A FH links
Deployment scenario 1	
RAT	MBS: Sub-6 GHz (CF: 3.5 GHz, BW: 100 MHz)
	SC: Sub-6 GHz (CF: 3.5 GHz, BW: 100 MHz)
Data requirement for Split-A	95.8 Gbps (MBS and SC)
Data requirement for Split-B	34.9 Gbps (MBS and SC)
Data requirement for Split-C	MBS: 34.9 Gbps with 100% of traffic load
	SC: 17.45 Gbps with 50% of traffic load
Data requirement for Split-D	MBS: 16.5 Gbps with 100% of traffic load
	SC: 8.25 Gbps with 50% of traffic load
Deployment scenario 2	
RAT	MBS: Sub-6 GHz (CF: 3.5 GHz, BW: 100 MHz)
	SC: mmWave (CF: 25 GHz, BW: 1 GHz)
Data requirement for Split-A	MBS: 95.8 Gbps
	SC: 574.56 Gbps
Data requirement for Split-B	MBS: 34.9 Gbps
	SC: 199.18 Gbps
Data requirement for Split-C	MBS: 34.9 Gbps with 100% of traffic load
	SC: 99.59 Gbps with 50% of traffic load
Data requirement for Split-D	MBS: 16.5 Gbps with 100% of traffic load
	SC: 42.32 Gbps with 50% of traffic load

loaded) and 200 SCs (50% loaded), which are connected to a common FH aggregator/BBU.

Figure 4.4 illustrates the basic idea of the considered scenario. Also, as mentioned in [64], the future wireless-based FH networks will be capacity constrained. Thus, in this optimization problem, we seek to find the optimal combination of different splits assuming the FH aggregator/BBU capacity is limited to a certain percentage of actual CRAN capacity requirements, i.e., we use as a reference the capacity required when all the FH links operate in Split-A (capacity required for Split-A x Number of FH links); that is, in the following, 100% capacity means that the network supports Split-A in all links. As an objective function, we consider two different approaches: (1) minimizing the OPEX, where finding the combination of different splits resulting into minimum cost is the objective, and (2) maximizing the centralization, where maximizing the number of FH links operating in Split-A is prioritized, and subsequently, priority decreases for the higher splits. We translated the presented scenarios into a Matlab code and utilizing a brute force algorithm, we find the smallest OPEX(€/year) among all possible combinations of splits.

We present the benefits of both the approaches in two different deployment scenarios; i) Deployment scenario 1 (cf. Table 4.3) considers all the MBSs and SCs operating in Sub-6 GHz band, and ii) Deployment scenario 2 (cf. Table 4.3) considers all the MBSs operate in the Sub-6 GHz band and the SCs in mmWave band, as suggested in [2]. Table 4.3 summarizes the parameters used for the performed evaluation. The presented FH requirements are calculated utilizing the equations presented in Section 2.4.1.

4.2.2 Evaluation results

In Deployment scenario 1, considering the first approach in which the overall OPEX is minimized, Figure 4.5 depicts the number of FH links operating in different splits for different levels of capacity availability in the FH aggregator/BBU. Utilizing a brute force algorithm (a deeper discussion on the application of a brute force algorithm is given in Section 4.3.2), we find the smallest OPEX(€/year) among all possible combinations of splits. As discussed in Section 4.1, each MBS contributes more to the operational cost and, therefore, the cost minimization strategy prioritizes Split-A links to MBSs over SCs. It is also beneficial in terms of the network performance, since MBSs are expected to serve twice the users of SCs [2] (100% load for MBSs vs. 50% load for SCs).

With increasing capacity availability in FH aggregator/BBU, the number of links operating with lower splits, i.e., Split-A/B, rises, and thus, higher centralization is achievable with minimum OPEX. Additionally, for capacity availability equal to or higher than 50%, all the links involved in the scenario belong to Split A and B, which allows almost full centralization

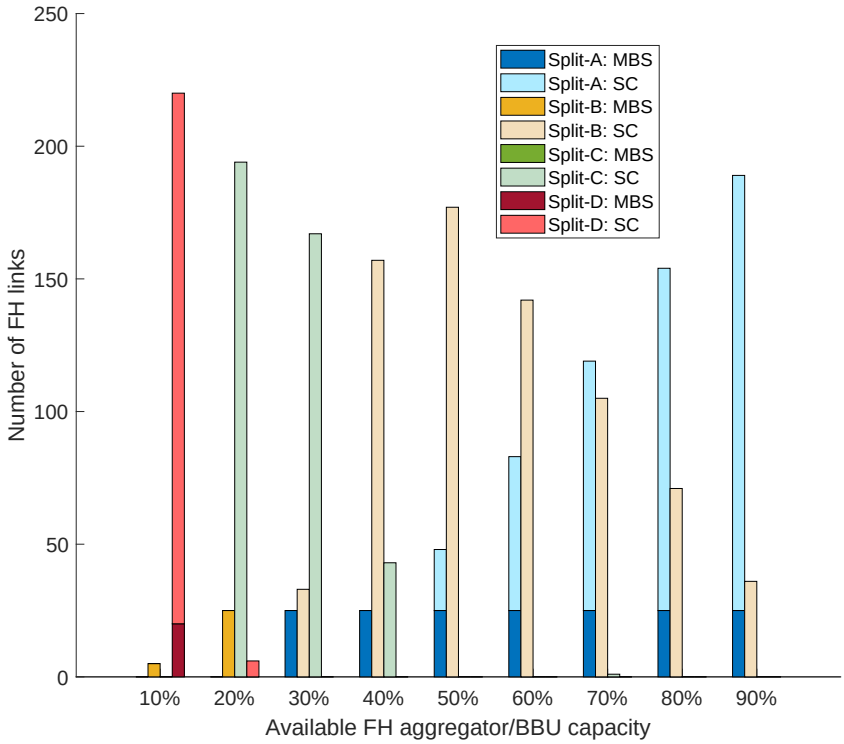

Figure 4.5: Scenario 1: Number of FH links utilizing different splits varying with different level of capacity availability in FH aggregator/BBU - minimizing OPEX approach.

gain, as discussed in Section 2.4.1 and Section 4.1. For congested scenarios, e.g., 10%, 20%, 30% of available capacity, it is still possible to serve all the FH links, even having some of them configured to Split-B, but higher functional splits at PHY layer are necessary, (i.e., Split C/D) to some extent.

Next, we focus on the second optimization approach in which the centralization is maximized. In this way, we find the best combination of links, prioritizing, in the first place, Split-A links for MBSs, and thereafter, Split-B, Split-C and Split-D links for MBSs, in the aforementioned order. Next, we follow the same procedure for SCs. Since MBSs will be serving a larger number of users (100% load for MBSs vs. 50% load for SCs), we prioritize MBS links in front of SC links.

Figure 4.6 illustrates the number of FH links operating in different splits for different levels of FH aggregator/BBU capacity, the second optimization approach, in deployment scenario 1. As expected, results show the same trend as the ones presented in Figure 4.5, since lower splits (i.e. Split-A/B) provide higher centralization, which also leads to lower OPEX. However, there are a few differences. For 20% of available capacity, when minimizing the OPEX, no FH link gets Split-A configuration, and Split-D is only considered for a few of

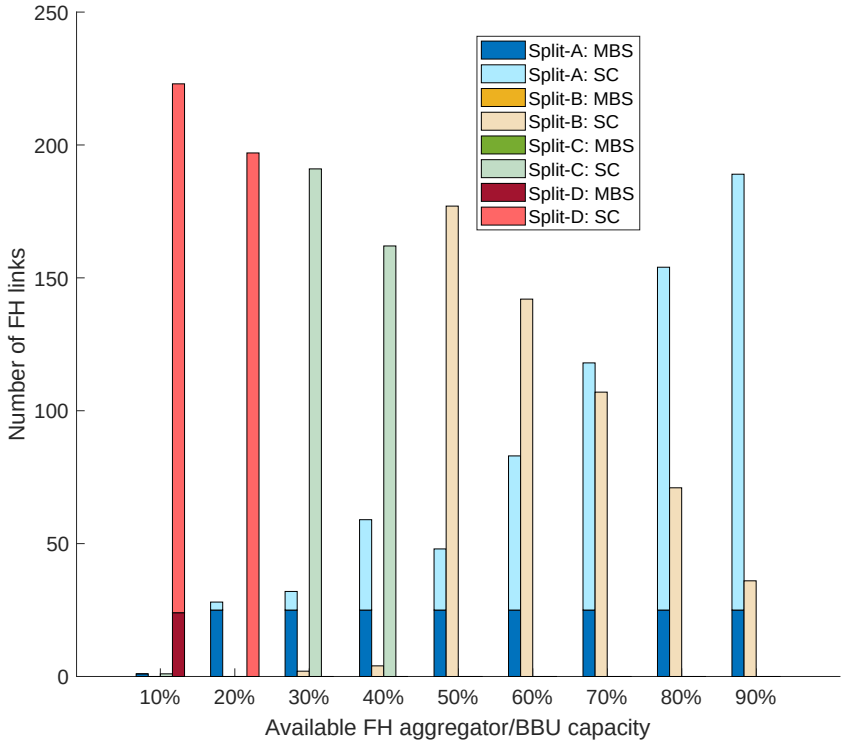

Figure 4.6: Scenario 1: Number of FH links utilizing different splits varying with different level of capacity availability in FH aggregator/BBU - maximizing centralization approach.

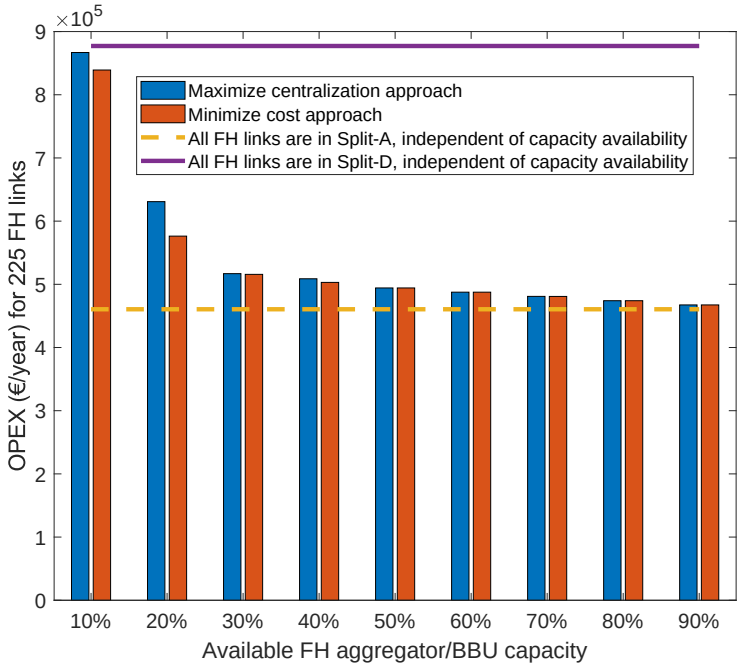

Figure 4.7: OPEX (€/year) for the heterogeneous split distributions illustrated in Figures 4.5 and 4.6

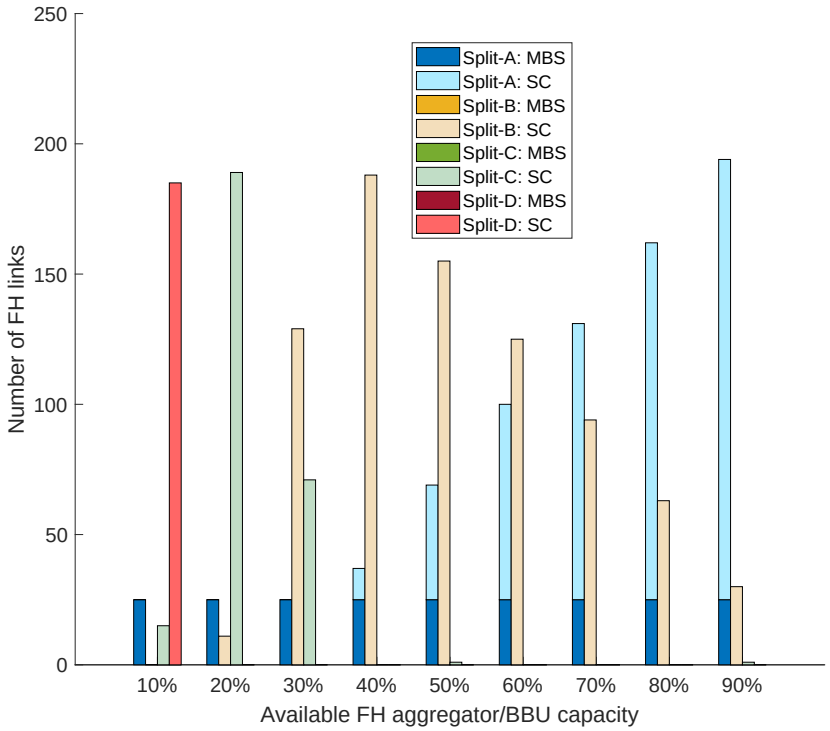

Figure 4.8: Scenario 2: Number of FH links utilizing different splits varying with different level of capacity availability in FH aggregator/BBU - minimizing OPEX approach.

them (Figure 4.5). On the other hand, under the same capacity conditions, and prioritizing centralization, all 25 MBS FH links operate with Split-A (Figure 4.6), thus achieving the highest centralization gain, and subsequently, a potentially better network performance for the MBS. However, remaining FH links, i.e. SCs links, have to sacrifice centralization and operate only with Split-D, thus increasing overall OPEX.

Figure 4.7 illustrates the total OPEX for 225 FH links and the configurations presented in Figures 4.5 and 4.6. For the discussed case, i.e. 20% of available capacity, the difference in cost is clearly visible. Hence, both approaches, i.e. minimizing OPEX and maximizing centralization, can lead to different combinations of splits at FH links under some capacity conditions (higher congested scenarios). Mobile operators can decide on the deployment rule to follow based on their particular interests.

We also provide a similar evaluation in a heterogeneous (i.e., MBS operating on Sub-6 GHz RAT and SCs operating on mmWave RAT) scenario, named deployment scenario 2 in Table 4.3. In this scenario, we have 25 FH links with Sub-6 GHz and 200 FH links with mmWave. Results are depicted in Figures 4.8 and 4.9, and follow the same trend as for scenario 1 in Figures 4.5 and 4.6, respectively. In this scenario, all the 25 MBS FH links are capable of achieving full-centralization, i.e. can operate under Split-A, utilizing 10% of

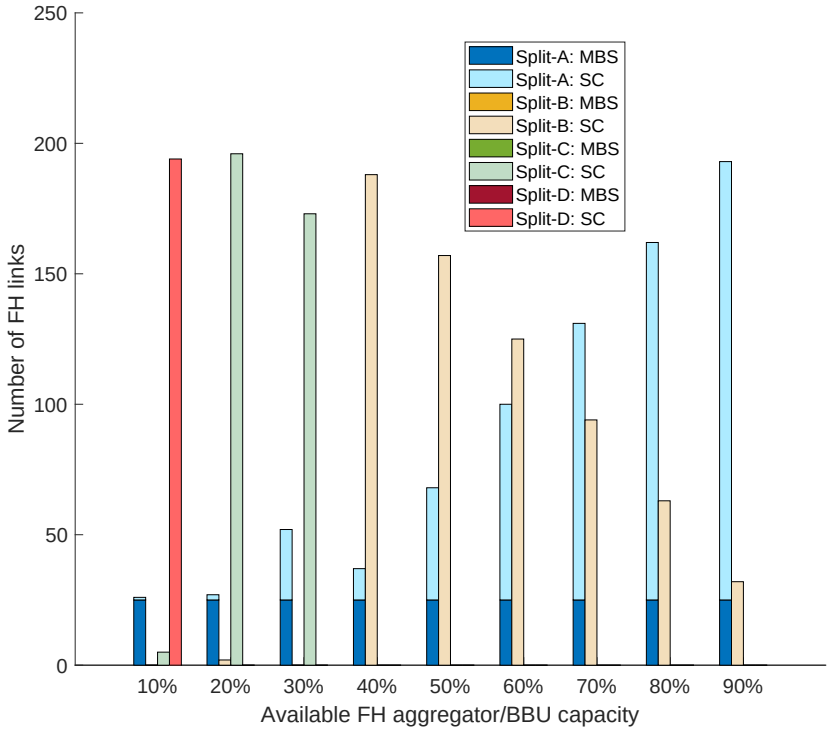

Figure 4.9: Scenario 2: Number of FH links utilizing different splits varying with different level of capacity availability in FH aggregator/BBU - maximizing centralization approach.

the required capacity at the FH aggregator/BBU. Note that, in this case, the maximum capacity considered (i.e., capacity required by using Split-A in all links) is larger than in scenario 1 due to the presence of demanding mmWave technology. Figure 4.10 depicts the OPEX comparison for the configurations presented in Figures 4.8 and 4.9, showing the same trend as for scenario 1.

Despite the different configurations of splits, both approaches show a similar cost when the available capacity is enough (i.e., 50% to 90%) to serve most of the MBS links at the lowest splits (i.e., Split-A). In this regard, the cost distribution shown in Figure 4.2 can be fine-tuned (e.g., weights of different PHY blocks can be slightly redistributed) to prioritize lower/higher splits following the operator's preference towards more centralized/distributed RAN[3].

Finally, we compute the FH Capacity Utilization Factor (FCUF), which we define as the ratio of **available** vs **utilized** capacity for different capacity conditions in FH aggregator/BBU. For both the deployment scenarios, i.e., all Sub-6 GHz and mixed Sub-6 GHz/mmWave, and all levels of capacity availability, i.e., 10%-90%, the FCUF is higher than 99.6%, for both optimization approaches. Therefore, the utilization of different splits

[3]Optimization of the cost function is out of the scope of this work.

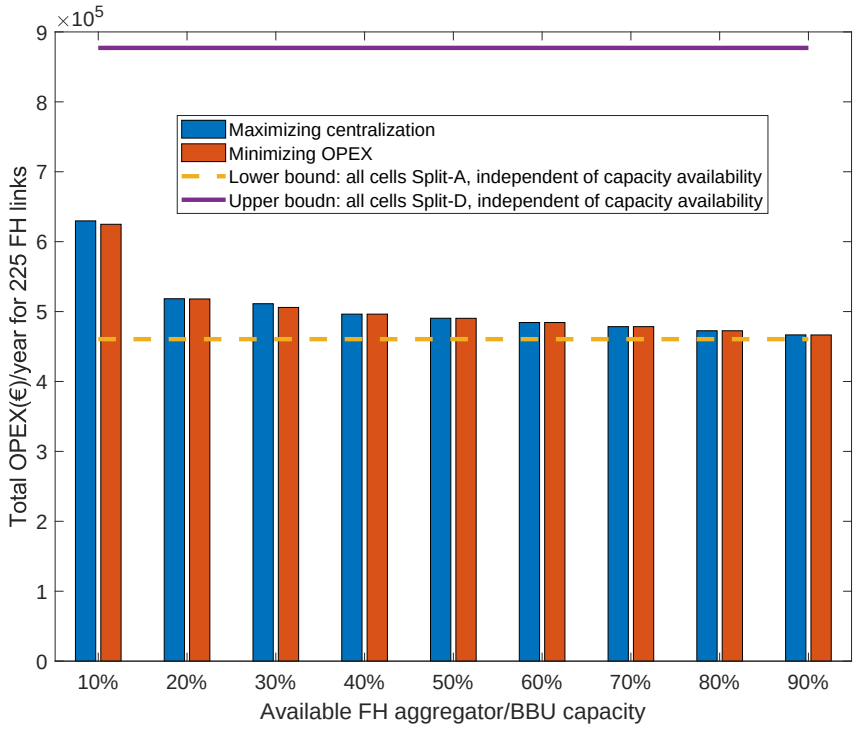

Figure 4.10: OPEX (€/year) for the heterogeneous split distributions illustrated in Figures 9 and 10.

at the PHY layer to generate an optimal combination for different congestion levels can ensure the best utilization of scarce resources.

4.2.3 Key takeaways

In the following, we summarize a few takeaways from the discussed results.

- Two potential RATs for 5G are presented in [C4] and [2], which are Sub-6 GHz (CF at 3.5 GHz with 100 MHz channel BW) for MBS, and mmWave (CF at 25 GHz with 1 GHz channel BW) for SCs.
- Capacity requirements for different splits depend on the RAT utilized in the access networks.
- Lower splits (i.e., Split-A/B) represent lower OPEX but are more demanding in terms of capacity and latency.
- To minimize OPEX, it is better to utilize lower splits, if the available capacity supports the higher capacity requirements.

- Variation of the cost for different splits is roughly ten times higher for MBSs than that for SCs. For instance, the OPEX difference between Split-C and Split-B in a MBS is ten times higher than the difference of Split-C and Split-B in a SC, although, the capacity requirements remain similar. This is because, according to Table 4.2, cost of different splits is directly proportional to the CAPEX of respective AP (i.e., MBS or SC), and MBS shows very high CAPEX. Thus, it is more cost-effective to choose lower splits for MBSs in front of SCs when the available capacity allows it. For SCs, higher splits (i.e., Split-C/D) can be utilized to remain within the available capacity.

- Utilizing a smart combination of splits, FH network can be deployed requiring only 10% of the total capacity (compared to pure Split-A option) at the FH aggregator.

- When the capacity is limited, employing higher splits (i.e., Split-C/D), especially in SCs, results in a fruitful solution to serve the FH links.

4.3 TCO-based optimization of fronthaul

The expected growth in carried traffic and the added complexity of different deployment approaches (distributed vs. centralized), showing different requirements, make the transport network one of the main design challenges in 5G. In a CRAN scenario, the anticipated capacity requirements for fronthaul links are enormous and, as of today, no single technology can support such requirements. There is consensus in the following two approaches to tackle the fronthaul challenge: i) building a heterogeneous network by combining different technologies; and ii) employing different functional splits, which have the potential to reduce the capacity requirements on fronthaul links, as discussed in Section 4.2. Hence, it is important that we exploit different potential technologies and a functional split approach for 5G fronthaul networks design. Advancing from Section 4.2, in this section, we show how intelligently selected functional splits at the PHY layer can be utilized to serve the radio access networks in a capacity-limited scenario, considering feasible FH technology options. From a different point of view, we also propose maximizing the centralization by means of a heterogeneous combination of functional splits in a TCO-limited scenario. Results presented in this section show that the combination of functional splits has the potential to enable the design of heterogeneous fronthaul networks combining wireless and wired links, and reducing drastically both the required capacity (to 42%) and the TCO (to 35%).

From the discussions in previous sections in this chapter, functional splits at the PHY layer can greatly relax the FH requirements, but at the cost of losing centralization. It is also discussed that the utilization of lower splits (Split-A/B), requiring higher data rates in the

FH, involves lower OPEX; on the other hand, higher splits (Split-C/D) require lower data rates, but entail higher OPEX. Therefore, as concluded in previous sections, a heterogeneous combination of splits will help in setting up the trade-off. Moreover, the utilization of different splits can be especially helpful when the FH aggregator (BBU or DU) is resource-limited. In such a resource-limited scenario, employing splits at different APs enable the FH to serve the RAN with a reasonable level of centralization, i.e., MAC layer or upper PHY layer functionalities are centralized. However, the above discussions and presented results are based only on a one-year OPEX forecast. At this point, it is also necessary to include the CAPEX and focus on TCO.

Table 4.4: Cost analysis of AP and BBU, the dependency on PHY layer splits [1].

Parameters	**CAPEX** (€)	**OPEX** (€/year)
Optical network central office	56,500	*10% of CAPEX*
Cost of fibre/meter (Urban area)	113.10 [72]	*1% of CAPEX*
E-band spectrum licensing fee	NA	*70.47* €/link [73]
Indoor equipment for wireless connection	7,830 [73]	NA

4.3.1 Analysis of the total cost of ownership

Now, we analyse the TCO of the FH network utilizing the cost assumptions presented in Table 4.2 and additional parameters presented in Table 4.4. We follow the previous section and cost distribution illustrated in Figure 4.2. Similar to Section 4.2, utilizing the cost distribution, cost (i.e., CAPEX or OPEX) of APs for different splits can be calculated as $AP_{Cost} \times f(S)$, where S refers to the split (A, B, C or D), and AP_{Cost} is the cost (i.e., CAPEX or OPEX) corresponding to an individual AP. Additionally, moving the splitting point towards or from AP, has impact on the cost associated to the BBU, which also varies for different splits as $BBU_{Cost} \times (1 - f(S))$. Note that, $f(S)$ represents the summation of cost functions W_s (i.e., weight of cost for different functionalities at PHY layer) presented in Figure 4.2, e.g., for Split-B, $f(B) = 0.4+0.13$.

We evaluate two different TCO approaches: i) TCO of capacity-limited RAN (following the approach in the previous section), and ii) TCO of RAN and cost of capacity at FH, where

the cost of capacity provided at the FH links is also considered. For both the approaches, we also present the optimal combination of heterogeneous splits at the PHY layer. Following the discussion on potential BH/FH options in Section 2.2, we consider two technologies, i.e., 100G fibre (100 Gbps) and mmWave (E-band, 51.2 Gbps) as the FH link options, ensuring the highest possible capacity available for wired and wireless technologies. As a secondary contribution, we also provide a study on candidate FH technologies, defining the related cost assumptions.

We follow the simple equations below to obtain the TCO of infrastructure in order to evaluate the first approach:

$$TCO_{\mathrm{N}} = CAPEX + N \times OPEX \tag{4.1}$$

$$CAPEX = CAPEX_{\mathrm{BBU}} \times (1 - f(S)) + CAPEX_{\mathrm{AP}} \times f(S) + CAPEX_{\mathrm{FHlink}} \tag{4.2}$$

$$OPEX = OPEX_{\mathrm{BBU}} \times (1 - f(S)) + OPEX_{\mathrm{AP}} \times f(S) + OPEX_{\mathrm{FHlink}} \tag{4.3}$$

Eq. 4.1 provides the TCO for N years, where CAPEX is the capital cost of the considered scenario, and OPEX is the operational cost. In Eq. 4.2, CAPEX of BBUs and APs are involved along with the corresponding summation of cost functions *f(S)*, as described earlier in Section 4.2, $\mathrm{CAPEX}_{\mathrm{FHlink}}$ is the CAPEX corresponding to the FH link between BBU and AP, which depends on the FH technology, i.e., fibre or wireless, and the additional costs related to it, e.g., optical network central office, wireless equipment, etc. (cf. Table 4.4). In the same way, Eq. 4.3 presents the OPEX.

Note that, in this study, we analyse the cost of the RAN, which we consider comprises the cost of all the elements between the BBU and the AP, both included. Costs beyond this segment of the network are out of the scope of this work. Additionally, the cost-benefit related to infrastructure sharing, as presented in [74] and [75], is also not considered in this work; rather, we consider a dedicated, non-shared network settlement.

However, it is important to include the cost of capacity in the FH links, and thus, Eq. 4.2 and Eq. 4.3 are turned into Eq. 4.4 and Eq. 4.5, respectively. Moreover, to evaluate the second approach, where TCO of RAN and cost of capacity at FH are considered, we introduce a cost factor, $\mathrm{C_f}$ (Eq. 4.6), which is the ratio between required capacity for a particular combination of splits vs. the maximum capacity at the FH aggregator, i.e.,

capacity required when all the FH links use the lowest split (more detailed discussion in Section 4.3.2). With this value, we discriminate between the cost of FH links requiring different capacity, assuming both CAPEX and OPEX increase linearly with capacity. In this way, a low value of C_f results into low CAPEX, OPEX, and subsequently, low TCO.

$$CAPEX = CAPEX_{\text{BBU}} \times (1 - f(S)) + CAPEX_{\text{AP}} \times f(S) + C_f \times CAPEX_{\text{FH link}} \tag{4.4}$$

$$OPEX = OPEX_{\text{BBU}} \times (1 - f(S)) + OPEX_{\text{AP}} \times f(S) + C_f \times OPEX_{\text{FH link}} \tag{4.5}$$

$$C_f = \frac{\text{Capacity required for a combination of splits}}{\text{Maximum capacity at the FH aggregator}} \tag{4.6}$$

The objective is to find the best combination of splits at the PHY layer while the system is assigned with limited capacity (different levels of capacity limitation at the FH aggregator (BBU or DU) were used). We used a brute force algorithm to select the functional split at each AP, which minimizes TCO as expressed in Eq. 4.1. A brute force algorithm searches all the possible solutions and selects the best result according to the objective function. The variables to be set in this TCO optimization problem (i.e., the split assigned to each AP) are discrete (split A, B, C or D) and, therefore, the solution space to the problem is finite (i.e., there is a finite number of possible split combinations). Considering the complexity of this scenario, a brute force algorithm, which guarantees that the optimal split distribution is always found after having evaluated all possible combinations, is deemed feasible and, in fact, provides a solution within a reasonable time period (~10 mins using a PC with Intel Core i9 at 3.30GHz and 62GB of RAM). The algorithm is further discussed in Section 4.3.2. For scenarios larger than those described in Section 4.3.2, and considering possible future extensions of the optimization function adding complexity to the problem, more efficient techniques may need to be explored, but this lies outside the scope of the present work. In scenarios similar (or less complex) to those studied in Section 4.3.2, such a brute force algorithm could be run on-demand or even periodically, thus enabling a dynamic approach. However, the study of the implications of dynamic splitting is left for future research.

In [69] and [76], a similar approach is proposed (i.e., cost-driven flexible functional split design), where the authors considered different levels of splits within layer 2. In [77], another cost model is presented, but its application is limited to the comparison between the two extremes CRAN and DRAN. As highlighted earlier, in this work we focus on the functional

splits within the PHY layer, which still allow a great level of centralization until the MAC layer (see Figure 4.2). Additionally, this work goes beyond the prior work (presented in Section 4.2) by evaluating the cost of respective FH technologies and capacity at the FH links, which we believe are essential to perform TCO-based analysis.

Furthermore, deployment challenges and some cost analysis of 5G CRAN have been discussed in the literature. In [78], authors investigated and proposed an optimization algorithm to minimize the use of fibre, thus reducing infrastructure costs, while finding an optimal placement for BBUs. Note that this work is not focused on BBU or site placement and, hence, we assume those parameters are already set. In [79], authors proposed an approach to minimize the cost of fronthaul network while also minimizing the cost of 5G cellular networks by exploring functional spits in CRAN. However, those works only consider optical fibre-based fronthaul. In our work, we explore both wireless and wired options and let the optimization algorithm have the freedom to choose the best FH technology and functional splitting according to the objective of minimizing cost on already set sites.

4.3.2 Description of the scenario

To evaluate the results utilizing various RATs, we consider the same two different deployments described in previous section (Section 4.2.1, Table 4.3). Respective feasible FH technology for different splits according to their requirements is mapped in Table 4.5.

Without loss of generality, we consider a large number of dummy scenarios, each of 1 km^2 area, served by 25 MBSs (100% loaded) and 200 SCs (50% loaded), where the BBU/DU/FH aggregator is located at the centre of the area. In each scenario, the MBSs are homogeneously placed in a grid fashion, but the SCs are deployed randomly over the whole area. In this way, many different FH link sizes are tested, being its average of, roughly, 350m. Among these 225 APs, in Scenario 1, all the SCs and MBSs operate utilizing Sub-6 GHz RAT, whereas, in Scenario 2, SCs utilize mmWave as RAT and MBSs operate with Sub-6 GHz. For both the scenarios, 100G fibre and mmWave E-band wireless options are considered as FH link technology.

Cost assumptions are derived from Table 4.2 and Table 4.4. We translated the described scenarios into a Matlab code and considered the system is capacity-limited. Before getting into the discussion of the results, in Table 4.5 we present the identified potential FH technologies (i.e., 100G and/or mmWave E-band) for different splits within both scenarios, and their corresponding costs and capacity are introduced in the algorithm.

We introduce two different modes of deployment, i.e., “fibre only”, where all the FH links are fibre-based, and “heterogeneous”, where different combinations of fibre and wireless

Table 4.5: Parameters used for evaluation

Scenario 1: MBS RAT: Sub-6 GHz, 100% loaded; and SC RAT: Sub-6 GHz, 50% loaded

Splits	**FH requirement (MBS)**	**FH technology**	**FH requirement (SC)**	**FH technology**
Split-A	95.76 Gbps	✓ 100G Fibre ✘ mmWave E-band	95.76 Gbps	✓ 100G Fibre ✘ mmWave E-band
Split-B	34.86 Gbps	✓ 100G Fibre ✓ mmWave E-band	34.86 Gbps	✓ 100G Fibre ✓ mmWave E-band
Split-C	34.86 Gbps	✓ 100G Fibre ✓ mmWave E-band	17.43 Gbps	✓ 100G Fibre ✓ mmWave E-band
Split-D	16.46 Gbps	✓ 100G Fibre ✓ mmWave E-band	8.23 Gbps	✓ 100G Fibre ✓ mmWave E-band

Scenario 2: MBS RAT: Sub-6 GHz, 100% loaded; and SC RAT: mmWave, 50% loaded

Splits	**MBS: FH requirement**	**FH technology**	**SC: FH requirement**	**FH technology**
Split-A	95.76 Gbps	✓ 100G Fibre ✘ mmWave E-band	574.57 Gbps	✘ 100G Fibre ✘ mmWave E-band
Split-B	34.86 Gbps	✓ 100G Fibre ✓ mmWave E-band	199.19 Gbps	✘ 100G Fibre ✘ mmWave E-band
Split-C	34.86 Gbps	✓ 100G Fibre ✓ mmWave E-band	99.60 Gbps	✓ 100G Fibre ✘ mmWave E-band
Split-D	16.46 Gbps	✓ 100G Fibre ✓ mmWave E-band	42.33 Gbps	✓ 100G Fibre ✓ mmWave E-band

technologies are used. We first evaluate the results with "fibre only" FH network, i.e., each AP requires a fibre-based FH link to connect itself to the FH aggregator (BBU or DU) and subsequently, we evaluate the "heterogeneous" case, where the algorithm has the freedom to select from fibre or wireless-based FH links according to the objective function. In order to establish a common reference for all the studied cases, we consider the maximum required capacity is the sum of capacity required in case all the Sub-6 GHz-based MBS are configured with Split-A, and all the mmWave-based SCs use Split-C. That is, when we vary the level of available capacity, it indicates the percentage with respect to the "Maximum capacity" = Number of MBS (25) × Split-A (RAT: Sub-6) requirement + Number of SC (200) × Split-C (RAT: mmWave) requirement, for all the upcoming set-ups. Note that, Split-A and Split-B for mmWave RAT cannot be supported by any of the considered FH technologies (Table 4.5) and, hence, we discard these two options.

As mentioned earlier, we used a brute force algorithm to try every possible combination. The algorithm presented in Algorithm 1 takes as arguments the total available fronthaul capacity (*AVACAPACITY*), total number of SC (*NUMSC*), total number of MBS (*NUMMBS*) and number of years to compute the TCO. As a result, the algorithm returns the number of SC in split C and D (SC_C and SC_D) [4], and the number of MBS in splits A, B, C, and D (MBS_A, MBS_B, etc.) that minimize TCO. The required fronthaul capacity for each split (Cap_SC_C, Cap_SC_D, Cap_MBS_A, etc.) is considered constant, and respective values are explained in Table 4.5. *MaxTCO* is also defined as a constant, taking an arbitrarily large value. The constraint of this problem is set in line 13 of Algorithm 1. According to the constraint, the required capacity (*reqCapacity*) for a particular combination cannot exceed *AVACAPACITY*, which we vary to generate the results for different levels of available capacity, as mentioned in Section 4.3.1 and discussed in the subsequent sections. The algorithm uses the functions *getCAPEX()* and *getOPEX()*, which provide CAPEX and OPEX, respectively, for a given configuration of splits, as defined in Section 4.3.1.

4.3.3 Minimizing TCO of capacity-limited RAN

We start the analysis of Scenario 1 (MBS and SC RAT: Sub-6 GHz) and utilizing Eq. 4.1, Eq. 4.2 and Eq. 4.3, for the case of an all-fibre FH. Figure 4.11 depicts the optimal number and split types carried by fibre FH links for different levels of available capacity when the objective is set to minimize the TCO for 1 year (see Section 4.3.5 for a more detailed discussion in larger time horizons) for Scenario 1. The trends shown in Figure

[4]We explain algorithm 1 applied to Scenario 2, which is easily extended to Scenario 1 enabling all splits also for SC.

Algorithm 4.1 Minimize TCO in a Capacity limited scenario

```
Input:  AVCAPACITY, NUMSC, NUMMBS, YEARS
Output:  SC_C, SC_D, MBS_A, MBS_B, MBS_C, MBS_D
Constants:  Cap_SC_C, Cap_SC_D, Cap_MBS_A, Cap_MBS_B, Cap_MBS_C, Cap_MBS_D,
MaxTCO
MinCost ← MaxTCO
for numSC_C = 0:NUMSC do
    numSC_D ← NUMSC - numSC_C
    for numMBS_A = 0:NUMMBS do
        for numMBS_B = 0:(NUMMBS - numMBS_A) do
            for numMBS_C = 0:(NUMMBS - numMBS_A - numMBS_B)  do
                numMBS_C = 0:(NUMMBS - numMBS_A - numMBS_B - numMBS_C)
                reqCapacity = numSC_C*Cap_SC_C + numSC_D*Cap_SC_D +
numMBS_A*Cap_MBS_A
                + numMBS_B*Cap_MBS_B + numMBS_C*Cap_MBS_C + numMBS_D*Cap_MBS_D
                if reqCapacity <= AVACAPACITY then
                    CAPEX ← getCAPEX(numSC_C, numSC_D, numMBS_A, numMBS_B, numMBS_C,
                    numMBS_D)
                    OPEX ← getOPEX(numSC_C, numSC_D, numMBS_A, numMBS_B, numMBS_C,
                    numMBS_D)
                    TCO ← CAPEX + YEARS * OPEX
                    if TCO < minCost then
                        minCost ← TCO
                        SC_C ← numSC_C
                        SC_D ← numSC_D
                        MBS_A ← numMBS_A
                        MBS_B ← numMBS_B
                        MBS_C ← numMBS_C
                        MBS_D ← numMBS_D
Return:  SC_C, SC_D, MBS_A, MBS_B, MBS_C, MBS_D
```

4.11 follow those that were observed in Section 4.2. The number of lower splits (Split-A/B) increases to minimize the cost when the available capacity allows it. Since regardless of the split chosen, the functionalities (and their associated costs) have to be installed either in the AP or BBU, the total CAPEX of AP and BBU remains the same and does not dominate the TCO-minimization objective. Additionally, for this "fibre only" case, the link cost is also not the dominating element since it is the same for all splits, rather; OPEX of the APs is the driving factor, i.e., the higher the available capacity, the larger the number of lower splits (Split-A/B) (Figure 4.11). Additionally, the optimal solution always prioritizes the lower splits for MBSs in front of SCs, because, as explained in Section 4.2.3, costs are notably higher for MBSs, as compared to SCs, even though provided capacity and PHY layer characteristics are the same. Note that, unlike Section 4.2, in this section, we have discarded two split options (i.e., Split-A and Split-B for mmWave RAT), which cannot be supported by the presented FH technologies (cf. Table 4.5.). Therefore, we recalculated the maximum capacity (cf. Section 4.3.2), and hence, as shown in Figure 4.11, the required minimum capacity at the FH aggregator is now 42%, which was 10% in Section 4.2.

Following, we introduce wireless technology at the FH network to observe the results for "heterogeneous" deployments. Figure 4.12 depicts the obtained results for Scenario 1, where both fibre and wireless options are available as FH technology. From Figure 4.12, it is clear that every time the algorithm has an option to select wireless FH technology, i.e., wireless FH-based Split-B (W) for both MBS and SC layer, it chooses the wireless option to minimize the TCO, since, according to our assumptions in Table 4.2 and Table 4.4, the wired option is always more expensive. It is also evident that, even with a higher capacity availability, it is cost efficient to avoid using Split-A, since corresponding data rate requirement can only be met by fibre-based FH (Split-A (F)), which is more expensive than wireless options, although Split-A presents lower OPEX. Thus, in this case, the dominant element for minimizing the TCO is the cost of FH link. Additionally, as the secondary dominant element for minimizing TCO, OPEX of the APs remains as in the previous case, and, thus, it is more cost-efficient to select the lowest possible split supported by wireless technology, which is Split-B, in this case.

Figure 4.13 presents the cost analysis for both the deployment modes, i.e., "fibre only" and "heterogeneous", in Scenario 1. CAPEX_F, OPEX_F and TCO_F represent the cost related to the "fibre only" mode, whereas CAPEX_F/W, OPEX_F/W and TCO_F/W represent the respective cost for the "heterogeneous" mode. Evidently, TCO for the "heterogeneous" deployment mode is less than the TCO for the "fibre only" deployment mode. What is more, "heterogeneous" deployment mode requires less TCO, i.e., TCO_F/W, than the CAPEX for the "fibre only" mode, i.e., CAPEX_F, if we take into account one-year costs. However,

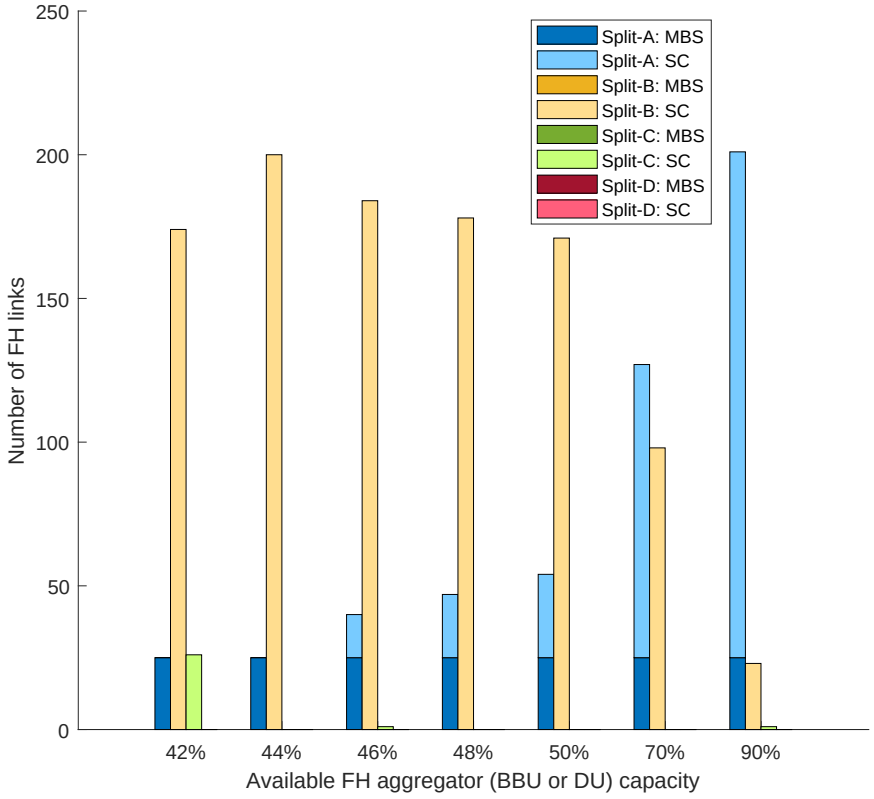

Figure 4.11: Scenario 1, "fibre only" mode: Number of FH links utilizing different splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

OPEX_F/W is slightly higher than OPEX_F; this is because the "fibre only" mode selects Split-A, whenever the available capacity allows it, and thus, it results into lower OPEX over the year. Additionally, OPEX_F decreases gradually as the number of lower splits (Split-A/B) increases. On the other hand, the combination for the "heterogeneous" deployment mode remains the same and, hence, no variation is experienced for OPEX_F/W.

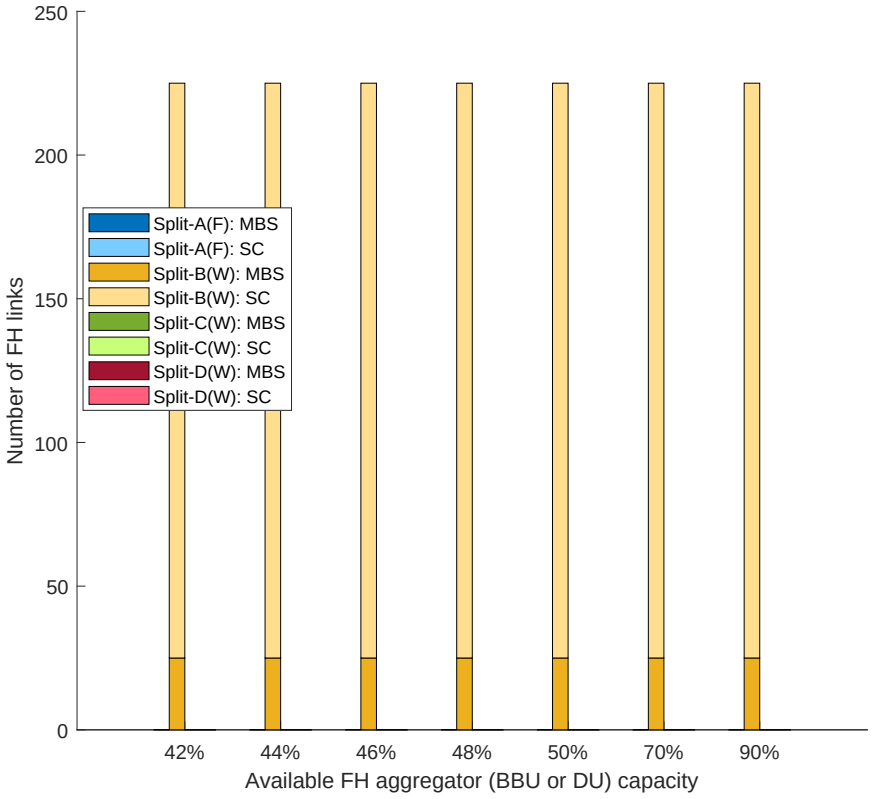

Figure 4.12: Scenario 1, "heterogeneous" mode: Number of FH links utilizing different splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

We perform the same evaluation for Scenario 2, where MBSs' RAT is Sub-6 GHz and SCs use mmWave. As depicted in Figure 4.14 for "fibre only" mode, with higher capacity availability, the combination with larger number of lower splits increases. As mentioned earlier in this section, Split-A and Split-B are not considered for the mmWave-based SC, hence, Split-C and Split-D are the only two valid options in this case. Thus, since lower splits require lower OPEX as found in Scenario 1 analysis, "fibre only" mode deployment is driven by OPEX of APs, and lower splits (Split-A for MBS and Split-C for SC) are selected to minimize the TCO. Once again, MBS Split-A has the higher priority because it reduces the required OPEX. On the other hand, Figure 4.15 shows the combination of splits at PHY layer for 225 FH links deployed in "heterogeneous" mode for Scenario 2. Since mmWave-based SC's Split-C cannot be supported by wireless technology due to very high capacity requirements (c.f. Table 3), Split-D is always selected for the SCs. On the other hand, Split-B is the lowest split for MBS, which can be supported by wireless options, and hence, Split-B is selected when enough capacity is available, i.e., from 44% at the FH aggregator

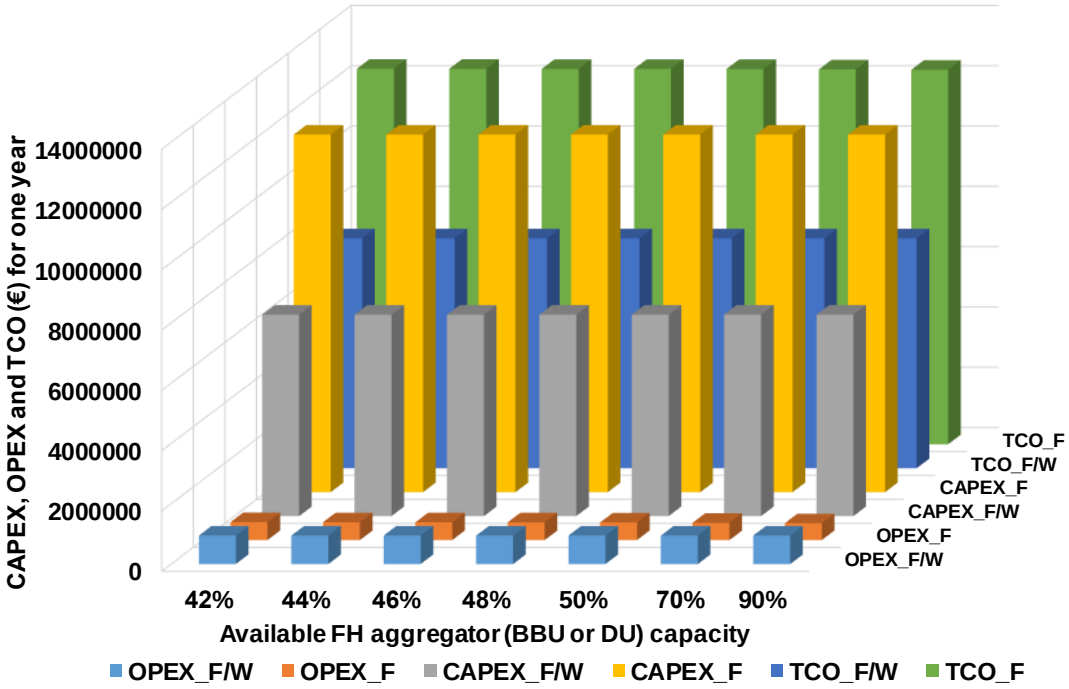

Figure 4.13: Scenario 1, Cost analysis for the combinations presented in Fig.4.11 and Fig.4.12.

(BBU or DU).

Cost analysis for both modes of deployment at Scenario 2 is presented in Figure 4.16, which shows similar behavior as the cost analysis presented in Figure 4.13. Similar to Scenario 1, in Scenario 2, TCO_F/W is lesser than CAPEX_F. OPEX_F/W is slightly higher than OPEX_F due to the "heterogeneous" mode not choosing the lower splits, and OPEX_F decreases with the more availability of lower splits.

Discussed results show that carefully selecting a combination of functional splits at the PHY layer, it is possible for an operator in capacity-limited deployments to design an efficient FH network and serve the RAN, with the minimum penalty in centralization while, at the same time, minimum costs of ownership are sought. It is also shown that although fibre-based FH network is more expensive, it validates the use of lower splits, hence increasing the centralization gain. On the other hand, when "heterogeneous" FH technologies are offered, minimization of the TCO leads to selecting the wireless options even using the higher splits, in front of fibre-based lower splits, which entail higher OPEX of the APs. In other words, it is more cost-efficient to select the lowest available splits supported by wireless technologies for FH links.

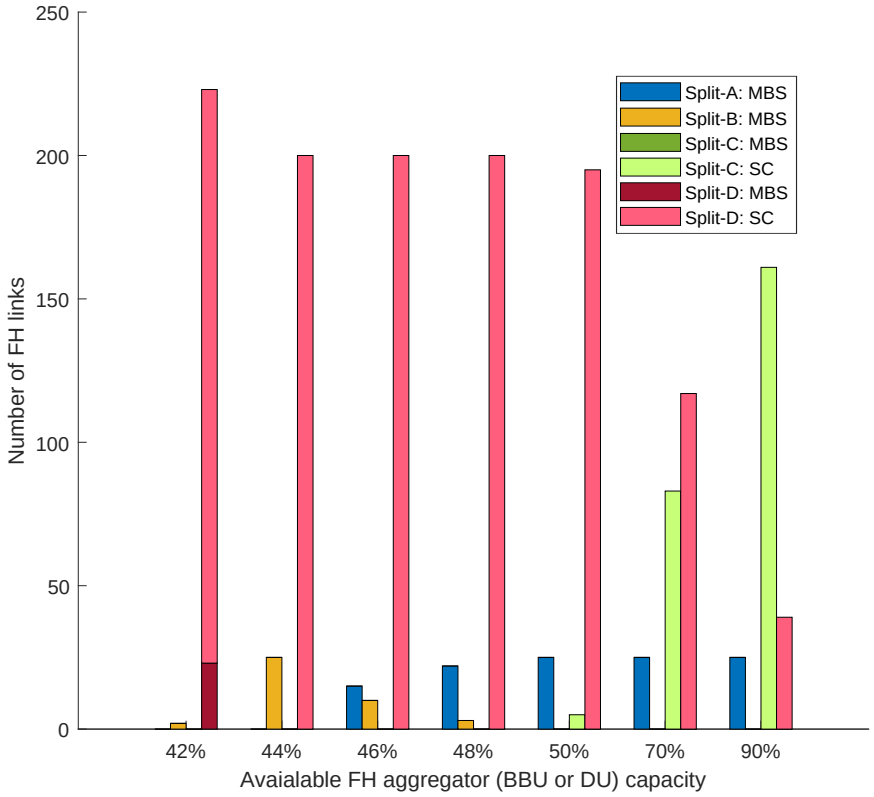

Figure 4.14: Scenario 2, "fibre only" mode: Number of FH links utilizing different splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

4.3.4 Considering the cost of capacity at the FH

The analysis presented so far considers the costs related to infrastructure, i.e., BBU, AP and additional equipment, spectrum licensing, etc, and assumes that the full capacity allowed by the technology is always granted at no cost, or as part of a prior (and already amortized) investment. However, other works in the literature suggest that this capacity should also be considered in the cost analysis [67], [14], [80]. Hence, to calculate TCO now we follow Eq. 4.4 and 4.5, as mentioned earlier in Section 4.3.1.

In the following, we assume a given capacity is provided at the FH aggregator, and we consider the effect that different capacity values assigned to the FH links should have on their cost by adding a "bias" penalizing capacity-hungry deployments.

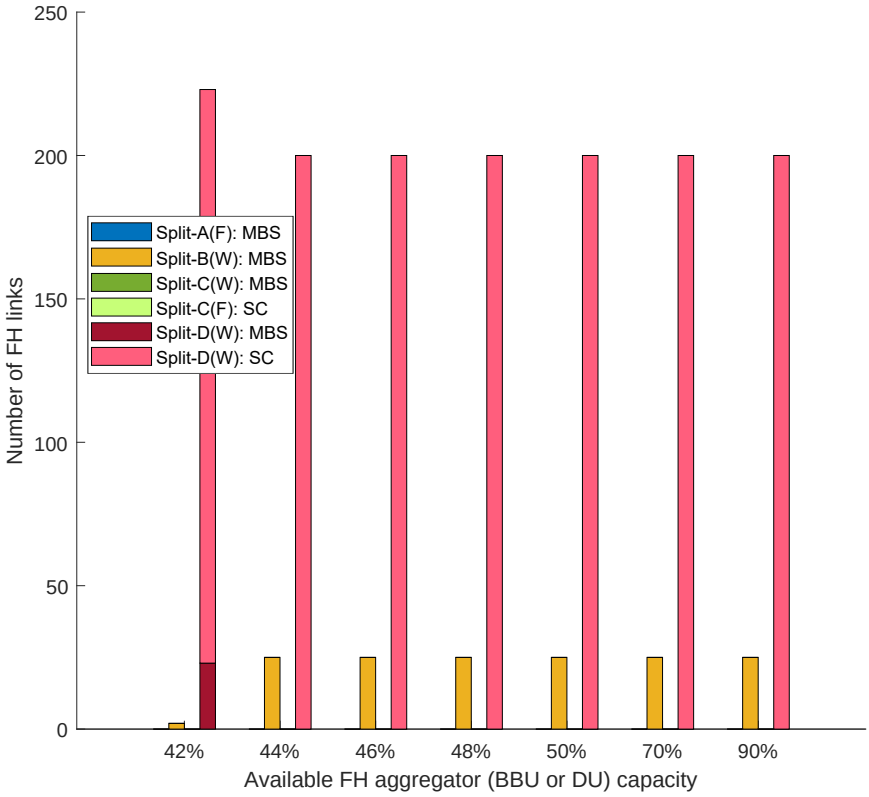

Figure 4.15: Scenario 2, "heterogeneous" mode: Number of FH links utilizing different splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

4.3.4.1 Evaluation of Scenario 1

Minimizing TCO in a capacity-limited scenario

Figure 4.17 presents the revised combination of splits at the PHY layer for Scenario 1 in the "fibre only" deployment mode. As mentioned earlier, now we include the cost of capacity in the same brute force algorithm (Algorithm 1) described in Section 4.3.2, and hence, larger capacity utilization results into higher OPEX, CAPEX and TCO. For this reason, minimizing TCO leads to selecting higher splits (Split-D). For the "fibre only" deployment mode, minimizing TCO was driven by OPEX in Section 4.3.3, in this case, it is driven by the cost of capacity at the FH links corresponding to SCs. On the other hand, for MBSs, the OPEX of APs is tipping the scale, hence, it is cost effective to select lower splits (Split-B), since higher splits (Split-D) ask for higher OPEX and the lowest split, (Split-A) increases the value of C_f. Additionally, it is also observed that the combination does not vary with the available capacity at the FH aggregator (BBU or DU). Even with the higher availability

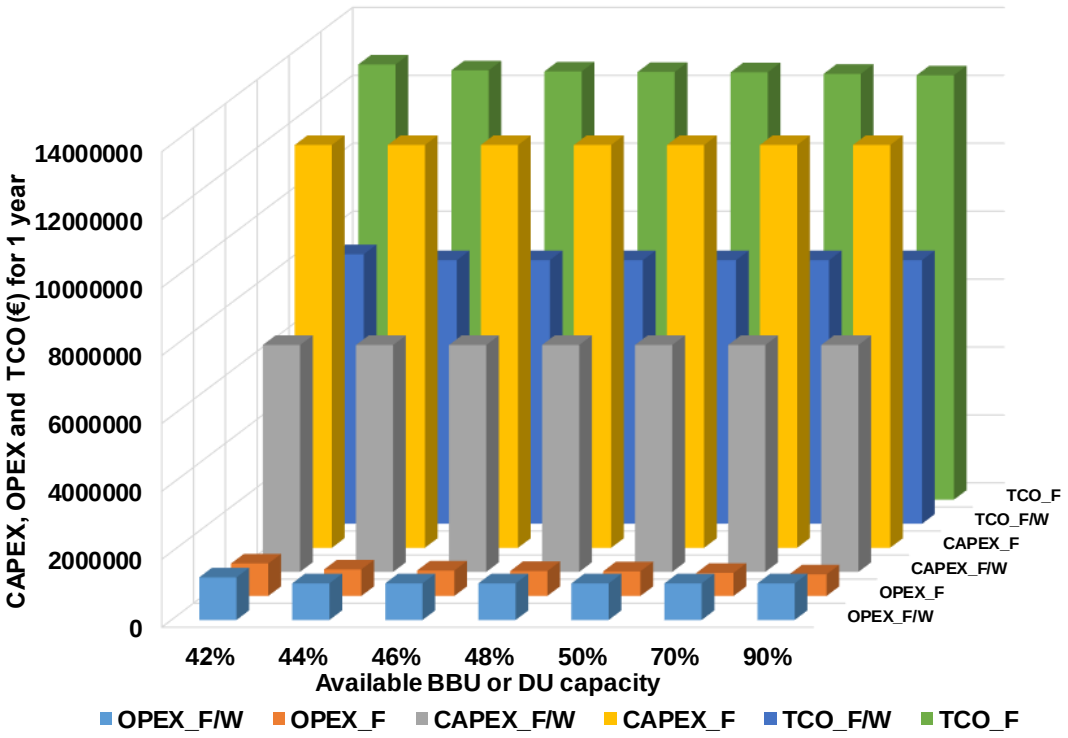

Figure 4.16: Scenario 2, Cost analysis for the combinations presented in Fig. 4.14 and Fig. 4.15.

of capacity at the FH aggregator (BBU or DU), it is still cost-efficient to select the same combination of splits.

In Scenario 1, and using the "heterogeneous" deployment mode, the combinations of splits are exactly the same as the ones presented in Figure 4.17. The optimal combination of splits to minimize TCO presented in Figure 4.17 for the "fibre only" mode can be supported by wireless technology, and, thus, the optimal results remain the same for the "heterogeneous" deployment mode. However, the technology for the FH links is now wireless, i.e., mmWave (E-band).

Cost analysis for the split combinations presented in Figure 4.17 for both fiber and heterogeneous modes is depicted in Figure 4.18. According to the updated equations presented in Eq. 4.4 and 4.5, we penalize the utilization of higher capacity with cost factor C_f (Eq. 4.6), and thus, the resulted OPEX, CAPEX and TCO are presented in cost units. One interesting finding from Figure 4.18 is that OPEX for wireless links is slightly higher than the OPEX for fibre links, and thus, with the same combination of splits, OPEX_F/W is slightly larger than OPEX_F. On the other hand, the TCO for the "heterogeneous" de-

ployment mode is less than the TCO for the "fibre only" deployment mode because of the higher CAPEX related to fibre deployment. Another observation from these results is that the available capacity at the FH aggregator (BBU or DU) is underutilized to reduce costs, therefore opportunities for new centralization gains are missed.

To maximize the centralization, in the following we define another objective: maximizing FCUF, which is the ratio of capacity utilized by a particular combination of splits in the FH network and the available capacity at the FH aggregator.

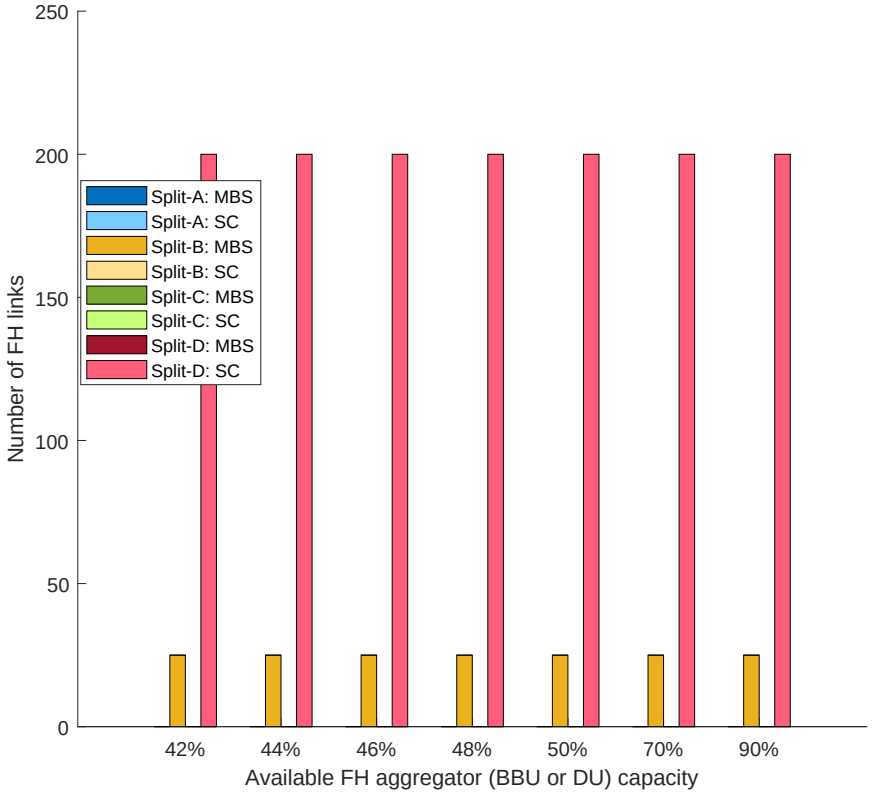

Figure 4.17: Scenario 1, "fibre only" mode: Revised combination of splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

Maximizing centralization in a TCO-limited scenario

At this point of the analysis, we adopt another approach to show the potential of the heterogeneous combination of splits at the PHY layer. Now we consider there are no capacity limitations at the FH aggregator (i.e., 100% of maximum capacity, which was defined earlier in Section 4.3.2, is available); rather the bottleneck is economical, i.e., limited TCO. Now we will find the best combination of splits for different levels of budget availability, where the objective is to maximize the centralization, i.e., maximize FCUF. Maximizing FCUF results

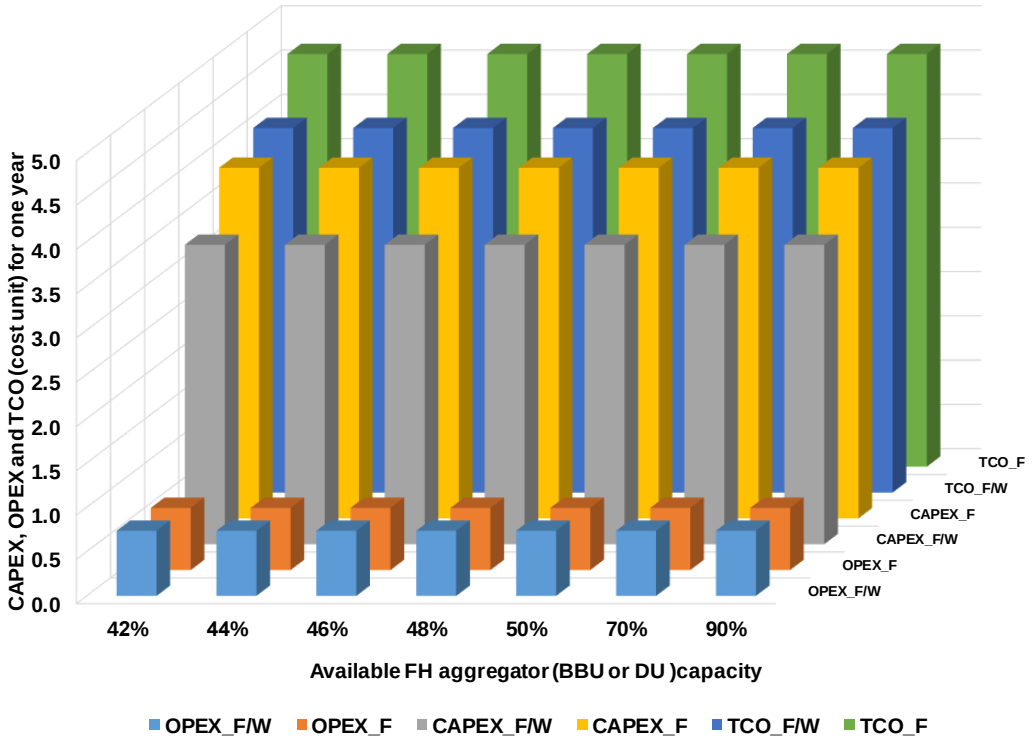

Figure 4.18: Scenario 1, Cost analysis of the optimal combinations for "fibre only" and "heterogeneous" mode - Minimizing TCO approach.

in a higher number of low splits (i.e., Split-A/B) in the FH links, benefiting the network with more centralization gains by allowing the implementation of advanced techniques such as CoMP, eICIC, etc. [C4]. This algorithm follows the same logic defined in Algorithm 1, changing the constraint-based on the available capacity ($AVACAPACITY$) for the maximum available TCO ($MaxTCO$).

In this analysis, lower splits (Split-A/B) are less expensive and provide higher centralization, on the contrary, recalling Eq. 4.4 and Eq. 4.5, there is an additional cost for the capacity provided at the FH links. Thus, the objective is to find the trade-off between OPEX, which is less for lower splits, and cost of capacity, which is higher for lower splits, while maximizing the centralization, which tends to increase the number of lower splits. Therefore, in this section, we show how to make the most of the available budget; it is the operator's decision when is the improved centralization worth the effort, given the increased TCO limit required.

The number of FH links varying with different TCO availability is presented in Figure 4.19 for Scenario 1 following "fibre only" deployment mode. Evidently, from the figure,

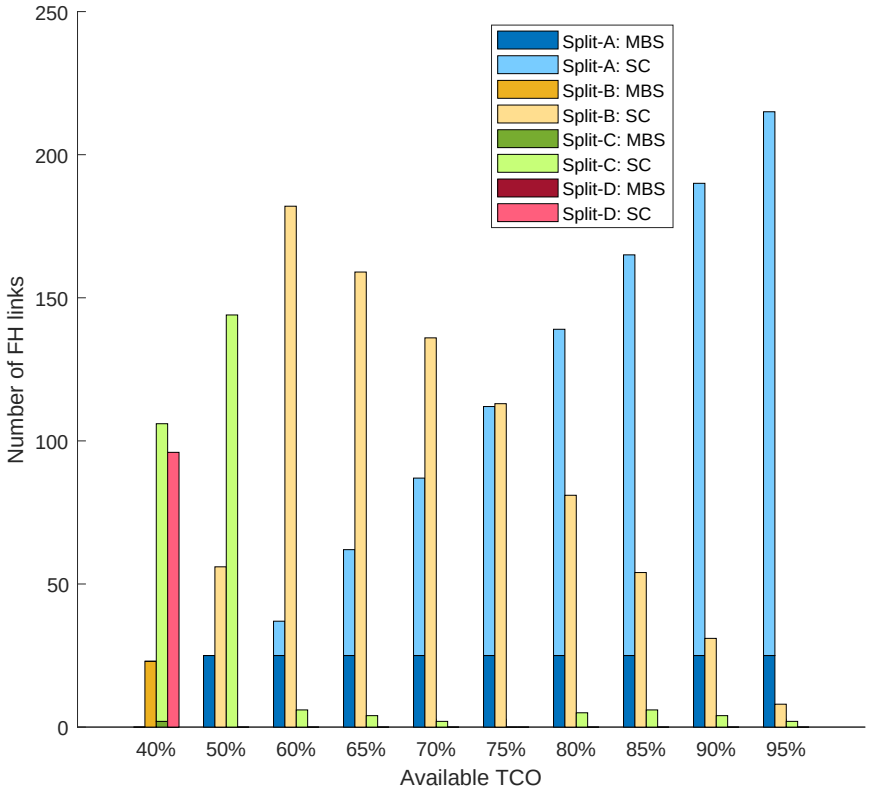

Figure 4.19: Scenario 1, "fibre only" mode: Number of FH links utilizing different splits varying with different level of TCO availability - Maximizing centralization approach.

utilizing only 40% of the maximum TCO, which is the TCO for the same combination we considered to limit the maximum capacity in Section 4.3.2, it is possible to deploy the FH network utilizing a heterogeneous combination of splits at the PHY layer. With the increment of the TCO availability, more links with lower splits (i.e., Split-A) are possible, which are usually more expensive, since, as mentioned in Table 4.5, Split-A requires higher capacity and fibre-based FH links. On the other hand, Split-A entails the lowest OPEX, provides more centralization benefits, and thus, better FCUF. For lower levels of TCO, the cost of capacity at the FH links is the driving element, and higher splits (Split-C/D) are selected. Additionally, it is observed that lower splits at MBS (i.e., Split-A/B) have a higher priority in front of those in SCs, due to the fact that lower splits for MBS lead to higher cost savings in terms of OPEX. Thus, when TCO is limited, it is more cost-effective to utilize lower splits (i.e., split-A/B) for MBSs than to do it for SCs.

Figure 4.20 shows the maximization of centralization approach with limited TCO for Scenario 1, utilizing "heterogeneous" deployment mode. Since we introduce wireless options in the FH network, FH deployment is possible utilizing only 35% of the maximum TCO (it

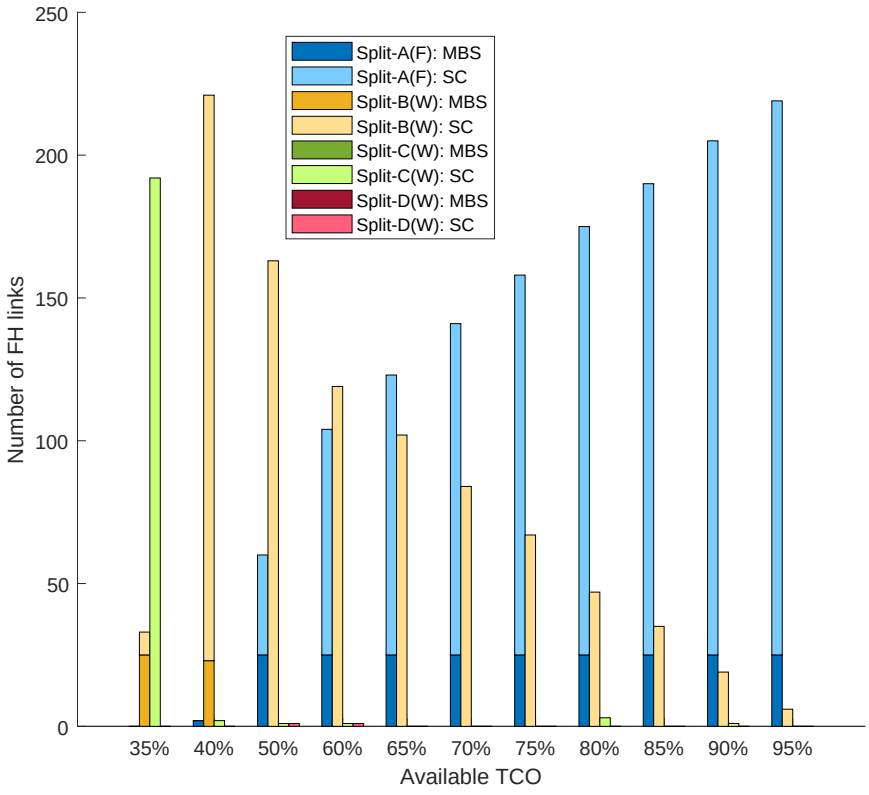

Figure 4.20: Scenario 1, "heterogeneous" mode: Number of FH links utilizing different splits varying with different level of TCO availability - Maximizing centralization approach.

was 40% for the "fibre only" mode). Additionally, after 50% to 60% of TCO availability, lower splits (Split-A/B) for MBS and SC are reached and Split-C is almost absent for the "heterogeneous" deployment mode, whereas it was always present (95% of maximum TCO) for the "fibre only" mode.

Figure 4.21 shows the cost analysis presented in cost units for this budget-limited maximizing centralization approach. Similar to the previous analysis, "heterogeneous" deployment mode results into less TCO requirements. Figure 4.22 shows the FCUF for both modes of deployment in Scenario 1. Evidently, utilizing the same level of TCO, "heterogeneous" deployment mode of FH networks provides higher centralization, i.e., it makes better use of available capacity at the FH aggregator.

4.3.4.2 Evaluation of Scenario 2

Minimizing TCO in a capacity-limited scenario

In the following, we provide the analysis with the revised TCO calculations utilizing Eq. 4.4 and Eq. 4.5 for Scenario 2. Figure 4.23 presents the revised combination of splits at the

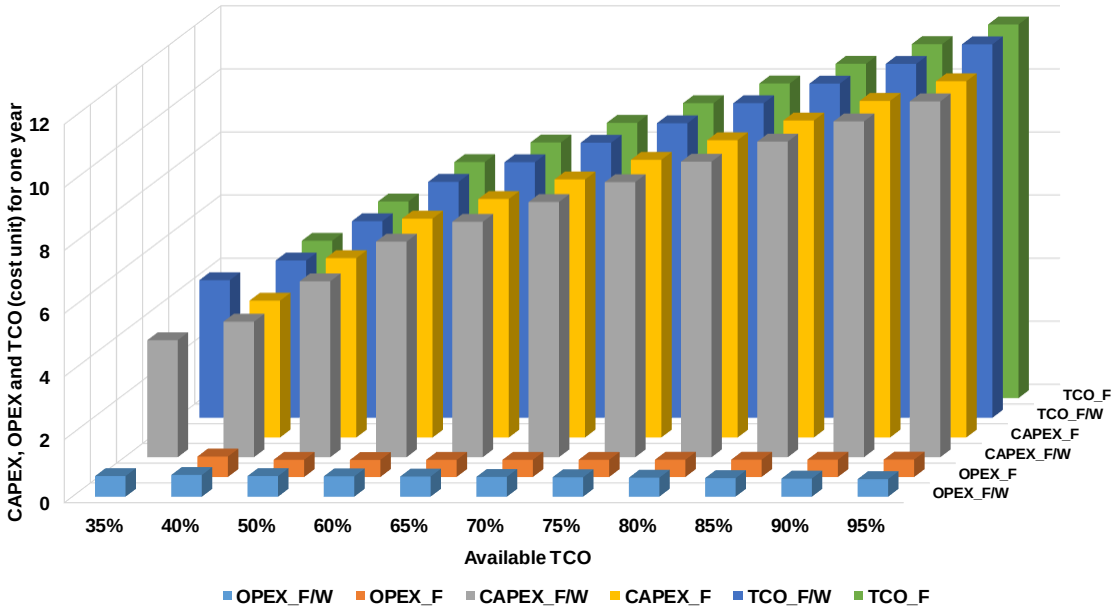

Figure 4.21: Scenario 1, Cost analysis for the combinations presented in Fig. 4.19 and Fig. 4.20 - Maximizing centralization approach.

PHY layer for Scenario 2 utilizing the "fibre only" deployment mode. Similar to Scenario 1 discussions, when the cost of capacity is introduced in the calculation, the TCO minimization approach for different capacity-limited scenarios selects higher splits (Split-D) for SCs, since higher splits require lower capacity in the FH links. For MBSs, similar to Scenario 1, Split-B is the optimal split, achieving a trade-off between OPEX and cost of capacity. Since the higher splits are selected even for the "fibre only" mode, which can be supported by wireless technologies, thus, the combination of splits remains the same for the "heterogeneous" mode. The cost difference is visible between these two approaches in Figure 4.24. Additionally, from Figure 4.24 it can be observed that, although the FH link combinations are similar (after 44% of available capacity) to the ones in Scenario 1 (Figures 4.17, 4.18), the cost (in cost units) is higher for Scenario 2 due to the fact that RAT for SCs is mmWave-based, and the corresponding FH links require higher capacity for the same combination of splits. Due to the additional cost of capacity, higher TCO is observed.

Maximizing centralization in a TCO-limited scenario

Similar to the *maximizing centralization* approach for scenario 1, below we discuss the analysis of maximizing the centralization in different budget-limited scenarios for both modes of deployment for Scenario 2. Figure 4.25 shows the number of FH links varying with different

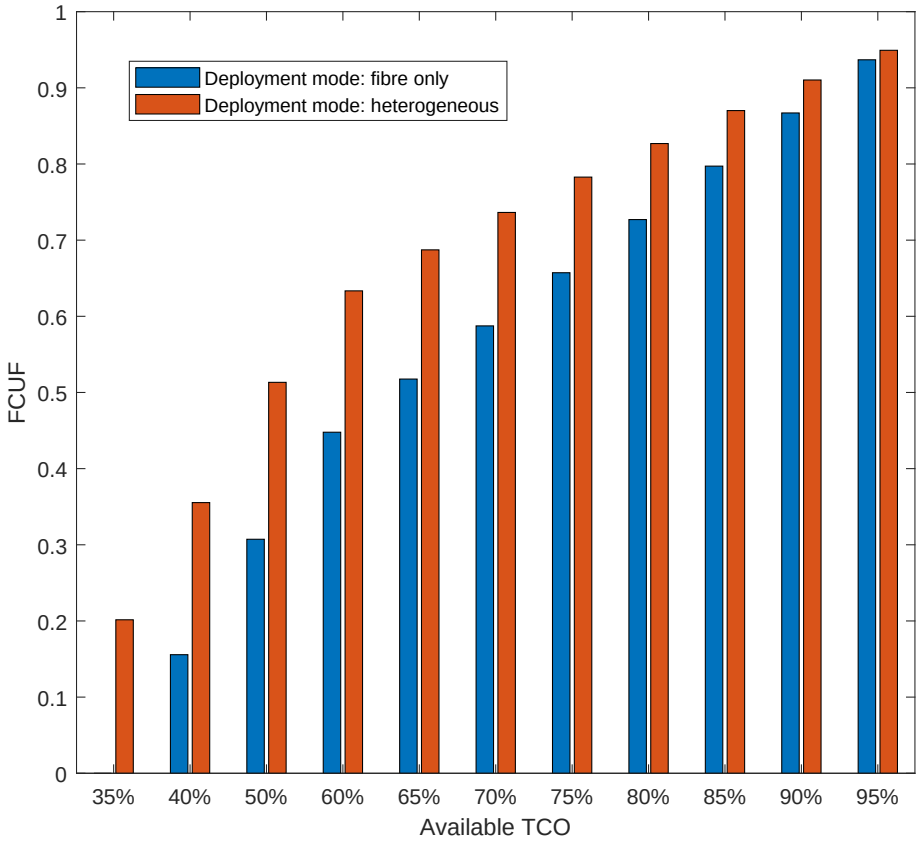

Figure 4.22: Scenario 1, FCUF comparison between the "fibre only" and the "heterogeneous" mode of deployments for the combinations presented in Fig. 4.20 and Fig. 4.21 - Maximizing centralization approach.

TCO levels for the "fibre only" deployment. Since the FH links for mmWave-based SCs require higher capacity in the FH links, the minimum TCO requires, at least, 60% of the reference TCO. Similar to earlier discussions, the number of the lower splits grows with the increasing availability of TCO. Figure 4.26 shows the same analysis for the "heterogeneous" deployment mode. Since wireless technologies are introduced, which are cheaper than fibre links, the minimum requirement of TCO is decreased to 45%. Figure 4.27 shows the cost analysis for both modes presented in Figure 4.25 and Figure 4.26. As expected, TCO of the "heterogeneous" deployment mode is always lower.

Figure 4.28 depicts the value of FCUF varying with the TCO level. Evidently, for the "heterogeneous" deployment mode, higher FCUF is achievable in comparison to the "fibre only" mode for the same level of TCO. Additionally, achievable FCUF for the "heterogeneous" mode utilizing 45% of the TCO is higher than that for the "fibre only" mode utilizing 60% of the TCO. Hence, a higher level of centralization is achievable with the "heterogeneous"

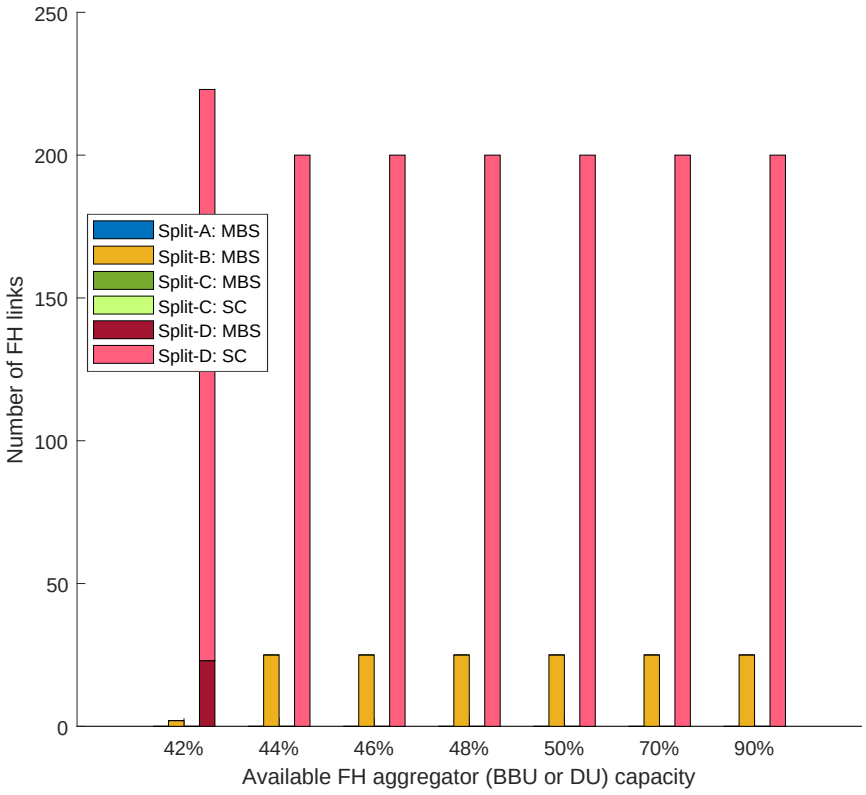

Figure 4.23: Scenario 2, "fibre only" mode: Revised number of FH links utilizing different splits varying with different level of capacity availability in BBU or DU - Minimizing TCO approach.

mode, utilizing the same or lower level of TCO, compared to the "fibre only" mode, due to the fact that wireless technologies are less expensive.

Discussed results in this section lead us again to conclude that, by utilizing an optimal combination of splits, an operator can deploy an efficient FH network and serve the RAN both in a capacity-limited or in a budget-limited scenario. Moreover, it is also observed that, from the operator's point of view, wireless FH options can be a very attractive solution since they are very cost-efficient, in spite of their capacity limitations. Note that, even in the "fibre only" mode, an optimization seeking a minimum TCO would select rather high splits (e.g., Split-B for MBSs and Split-D for SCs), while in Section 4.3.3, where the system was agnostic to the higher cost of capacity required by the lower splits, Split-A for MBSs and Split-C for SCs were preferred. On the other hand, seeking to improve centralization gains, lower splits are selected when available TCO allows to. Thus, an operator can have its priority fixed, i.e., minimize TCO or maximize centralization (or a trade-off between them), and deploy accordingly.

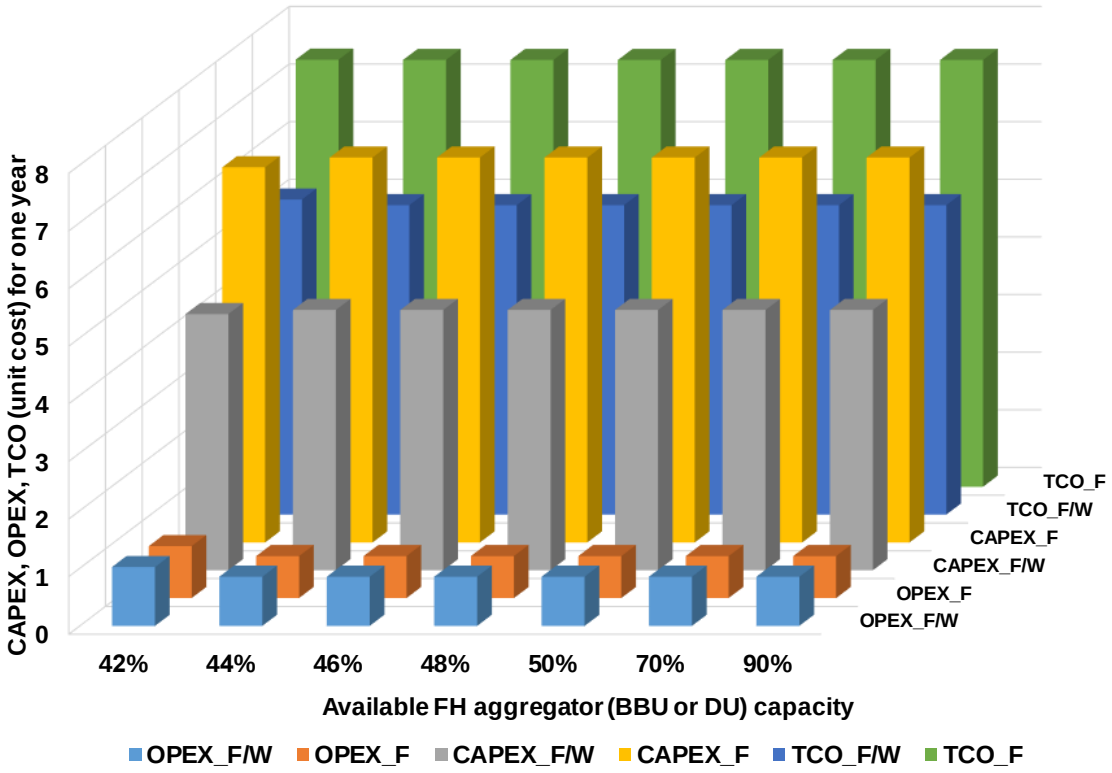

Figure 4.24: Scenario 2, Cost analysis of the optimal combinations for "fibre only" and "heterogeneous" mode - Minimizing TCO approach.

4.3.5 Additional findings

Results discussed in the previous sections consider 1 year TCO of a RAN using different functional splits within the PHY layer. We performed the same analysis for 5 and 10 years. For the TCO minimization approach, i.e., capacity-limited scenarios, the best combination of splits do not change for 5 or 10 years deployments considering both scenarios, i.e., Scenario 1 and 2, and both deployment modes, i.e., "fibre only" and "heterogeneous".

As shown in Figure 4.29, 5-year OPEX is closer to total CAPEX, especially for the "heterogeneous" deployment. For "fibre only" mode, the differences between CAPEX and 5-year OPEX are still significant. On the other hand, after 10 years, OPEX overtakes CAPEX. As explained earlier in Section 4.3.3 and 4.3.4, the combination of splits is mostly driven by the OPEX, and hence, the combination remains the same for 1, 5 and 10 years, given that OPEX becomes even more important when we increase the number of years considered.

Additionally, it is clear that, Split-B for MBSs and Split-D for SCs are generally preferred split options for minimizing TCO and it remains the same for longer periods of time (i.e., 10 years). Thus, although, the initial OPEX-based study [C4] suggested that lower splits

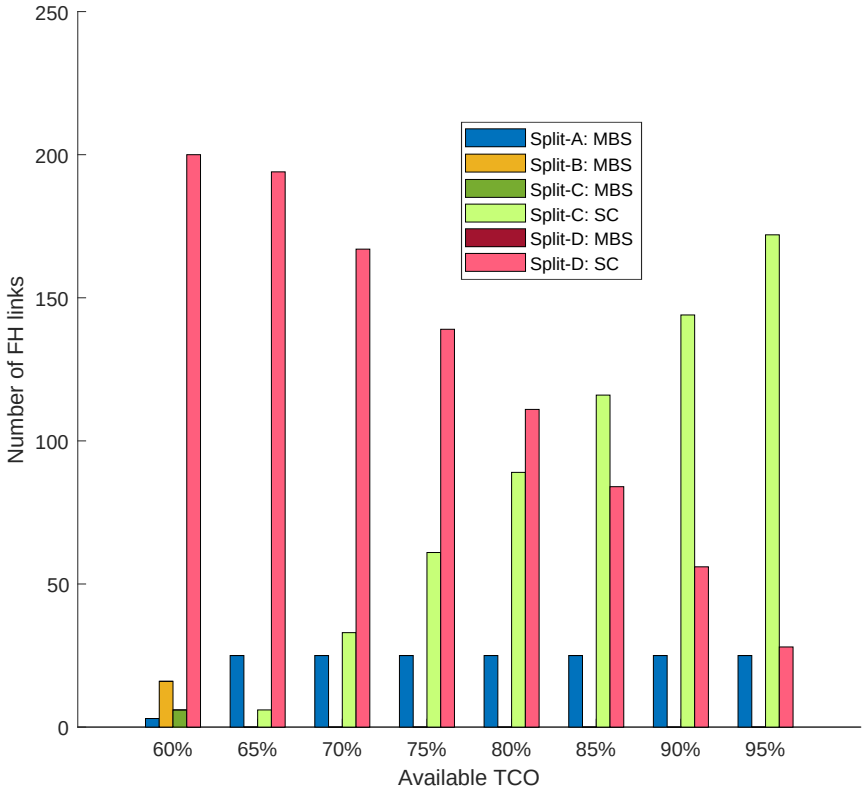

Figure 4.25: Scenario 2, "fibre only" mode: Number of FH links utilizing different splits varying with different level of TCO availability - Maximizing centralization approach.

(Split-A/B) are better for minimizing cost, that statement, however, needs to be revisited considering TCO and the cost of capacity provided in the FH links. Rather, in "heterogeneous" mode, Split-B for the expensive MBSs and Split-D for less expensive SCs are found to be the optimal splits, from an overall cost perspective, even with higher availability of capacity at the FH aggregator. To minimize cost under "fibre only" mode, combination of different splits at the PHY layer becomes useful to serve the RAN in a capacity-limited scenario.

We have also tested the centralization maximization in a TCO-limited scenario approach for 5 and 10 years. The combination follows the same trend as explained earlier in Section 4.3.2, 4.3.3, and 4.3.4. With this budget availability, higher centralization is possible, since lower splits (Split-A/B), requiring higher capacity, take advantage of the abundant resources. Moreover, utilizing the "heterogeneous" mode, it is possible to reach higher centralization than with the "fibre only" mode for the same, or even less TCO.

To discuss the sensitivity of the assumed cost values, we recall our budget, the maximum TCO, which is presented in Section 4.3.4 as the TCO of the combination to limit the maxi-

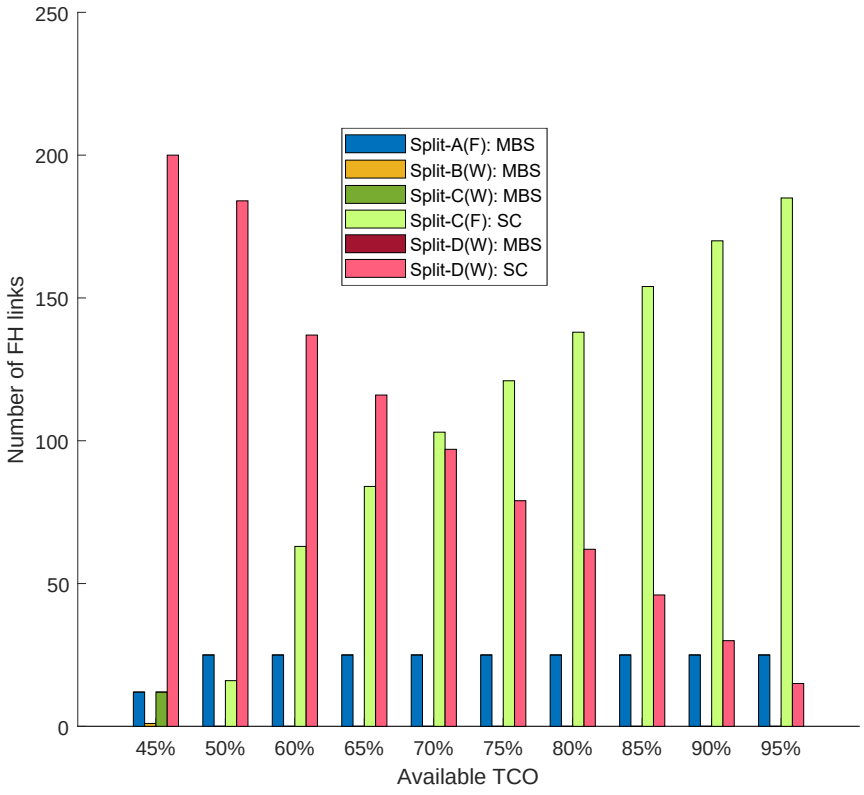

Figure 4.26: Scenario 2, "heterogeneous" mode: Number of FH links utilizing different splits varying with different level of TCO availability - Maximizing centralization approach.

mum capacity in Section 4.3.2. In this budget for 1 year, 95.08% of the budget belongs to the CAPEX, whereas, the rest belongs to the OPEX. On the other hand, 71.4% of the TCO and 75% of the total CAPEX belongs to the CAPEX of the deployed fibre links; that is, a 1% increase in the cost of fibre deployment is reflected in a 0.75% increase in the total CAPEX (0.7% increase in one-year TCO). Moreover, according to our cost assumptions, for a single link, wireless options showed a CAPEX more than 50% lower than that of fibre options. Therefore, every time the wireless options are suitable, the objective function prefers the wireless option in front of the fibre links to minimize the TCO.

4.4 Summary and future work

In this chapter, we have reviewed and analysed PHY layer functional splits to address one of the most anticipated challenges in 5G, presented in FH link design. A novel solution utilizing different combinations of splits was presented to tackle this challenge. With this solution, a trade-off between centralization and required capacity can be achieved, which, on the other

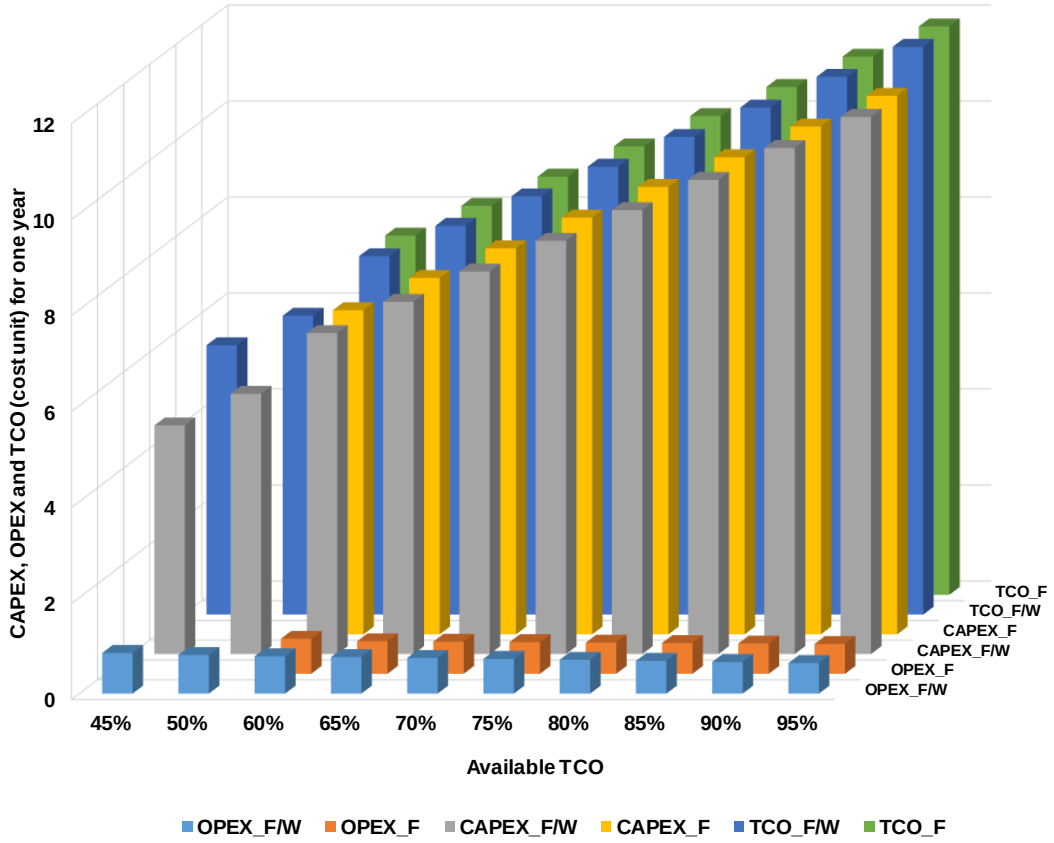

Figure 4.27: Cost analysis for the combinations presented in Fig. 4.25 and Fig. 4.26 - Maximizing centralization approach.

hand, has a major impact on OPEX (cf. Section 4.2), and thus, on TCO (cf. Section 4.3).

In Section 4.2, it is shown that, with the higher splits at the PHY layer (i.e. more decentralized approach), total OPEX (BBU and AP) increases. Therefore, we presented two techniques to maximize the centralization and minimize the OPEX utilizing different combinations of heterogeneous split deployments. We also presented the combinations for different levels of capacity availability in FH aggregator/BBU, showing how utilization of splits can still achieve centralization for high priority APs (e.g., MBSs), in capacity-limited scenarios (utilizing 10% of the required capacity). Additionally, OPEX comparison for different optimization approaches (i.e. minimum OPEX and maximum centralization) are discussed to evaluate the costs of maximizing the centralization.

In Section 4.3, the presented results show that, even if the FH aggregator is capacity-limited, most aggressive (i.e., closer to the antenna) splits can still be supported by the FH, at the cost of reducing centralization gains in other sites. To minimize TCO, Split-B for

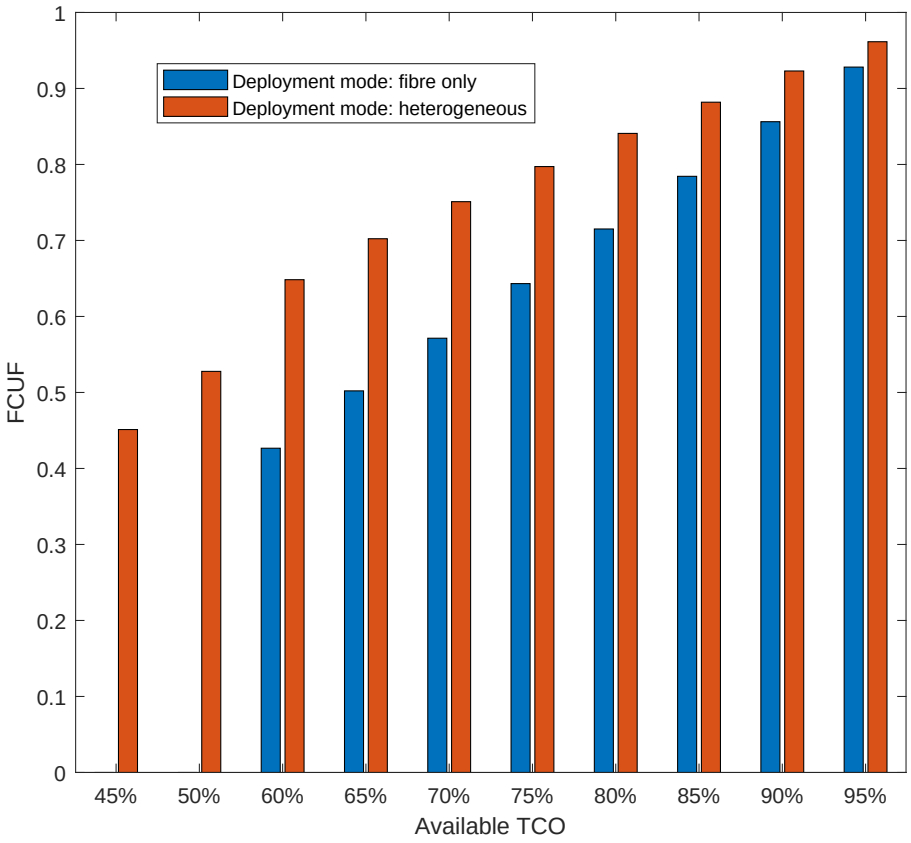

Figure 4.28: Scenario 2, FCUF comparison between "fibre only" and "heterogeneous" mode of deployments for the combinations presented in Fig. 4.26 and Fig. 4.27 - Maximizing centralization approach.

MBS and Split-D for SCs, were found to be the optimal choice for different RATs (i.e., Sub-6 GHz and mmWave). Additionally, to maximize the centralization, a different approach seeking combinations of splits to remain within a limited budget was also presented. Such combinations can be very useful to deploy the future FH networks with special care dealing with the trade-off between FH cost and centralization benefits. Comparative studies and cost analysis of "fibre only" and "heterogeneous" (i.e., mix of fibre and wireless-based links) FH deployment modes were also presented. Discussed results show the potential economic benefits (50% lower CAPEX) of wireless technologies used in FH links.

Finally, we add the cost of capacity to the analysis (e.g., with application in case the required capacity has to be leased). In this case, our analysis shows that a more conservative approach is preferred (i.e., use of higher splits) for TCO minimization.

Presented results showed that such combination of splits can be a very efficient solution for the cost (up to 35%) and capacity-limited (up to 42%) scenarios. Thus, for our future

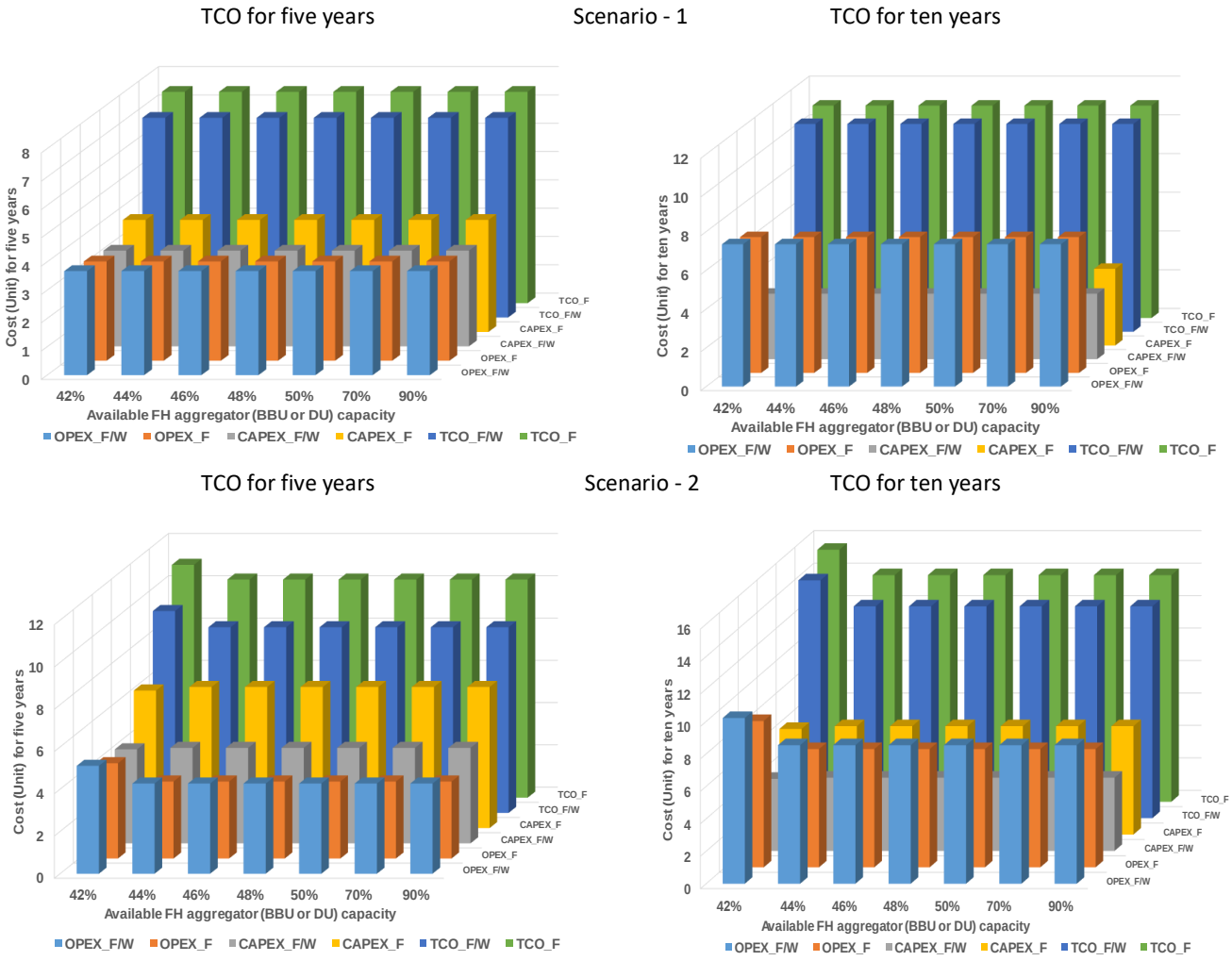

Figure 4.29: Different costs for 5 and 10 years of deployment - Scenario 1 and 2 - Minimizing TCO approach.

work, we plan to evaluate this approach in a larger scale, where a more intelligent algorithm is required. Another interesting extension of this work is the consideration of a dynamic approach where, with the help of an SDN-based controller, the choice of heterogeneous splits could be a real-time decision, based on the dynamic traffic demands of the network, which affects capacity requirements in deployments using higher splits. Subsequently, the adaptability of the equipment, related cost, the multiplexing gain, cost-benefit in a shared infrastructure type of scenario and the impact on the TCO are also worth studying.

Furthermore, splits throughout the upper layers, i.e., layer 2 and their impact on the TCO, are also interesting for future studies.

Chapter 5

Joint spectrum sharing

Every generation of cellular networks, i.e., 1st, 2nd, 3rd and 4th Generation improved the performance of the previous generation in terms of capacity, coverage, data rate, etc, to meet the expectations of an evolving and connected society. Moreover, with the increasing popularity of mobile devices, tactile Internet applications, video streaming and multi-fold varieties of use cases (e.g., broadband access everywhere, higher user mobility, extreme real-time communications, ultra-reliable communications), wireless networks are becoming more popular as a cost-effective solution for ubiquitous connectivity. In that sense, 5G is considered revolutionary, since it is promising a level of services and facing new challenges like never before.

To meet the expected demands from 5G (high capacity, low latency, etc.), provide ubiquitous coverage, and be able to serve different scenarios (IoT, D2D, Vehicle to everything (V2X), enhanced mobile broadband (eMBB) communications, etc.), wireless networks are becoming more complex everyday. Disrupting features of future wireless networks, such as *MIMO*, *Multi-RATs*, *CRAN*, *Network slicing*, *Edge computing*, *NFV*, etc. are the key enablers of 5G to meet the promised level of QoS. Thus, implementing highly dense, complex and multi-layered 5G networks will require a higher degree of automation [81]. Although, according to [82], the existing 5G networks do not provide such level of flexibility or automation yet, Artificial Intelligence (AI) offer solutions to tackle the complexity of 5G and beyond [81] [82]. Among many other AI techniques, (e.g., autonomous vehicles, robotics, computer vision, etc.), Machine Learning (ML) is probably the most convenient mechanism since it depends on the availability of large amounts of data, something that abounds in a modern mobile network.

On the other hand, cell densification has always been the most effective and fastest way to increase the area capacity of the network, however, this process is costly and introduces addi-

tional challenges such as interference management. Deployment of small cells and techniques like eICIC are popular solutions for the aforementioned cost and interference challenges, respectively. Largely deployed SCs along with network-wide deployed IoT devices will result into an ultra dense network [C3]. The aggregated capacity requirement of such UDN will be enormous and potentially unbearable for current capacity enhancement techniques, such as carrier aggregation or Frequency Reuse (FR), because of poor frequency resource management: while some coverage areas are overprovisioned, others appear overloaded. In this way, frequency resources are becoming one of the most valuable assets, which require proper utilization and fair distribution. Traditional frequency resource management strategies are often based on static approaches, and are agnostic to the instantaneous demand of the network. These static approaches tend to cause congestion in a few cells, whereas at the same time, might waste those precious resources on others. Therefore, such static approaches are not efficient enough to deal with the capacity requirement challenge of the future network.

Hence, in this chapter, we propose and evaluate a novel technique to share the available spectrum from a common spectrum pool, and test the proposed spectrum sharing technique following two different use cases for future wireless networks. First, in Section 5.1, we present the idea of spectrum sharing among different links from a cooperative access-backhaul mechanism point of view. We present simulation results for different approaches of such sharing from a common spectrum pool. The results show that traffic-aware approaches show increased fairness, thus reinforcing the idea of cooperative access-backhaul mechanisms as essential strategies in current and future networks. Second, in Section 5.2, we present a dynamic access-aware bandwidth allocation approach, which follows the dynamic traffic requirements of each cell and allocates the required bandwidth accordingly from a common spectrum pool, which gathers the entire system bandwidth. We perform the evaluation of our proposal by means of real network traffic traces. Evaluation results presented in Section 5.2 depict the performance gain of the proposed dynamic access-aware approach compared to two different traditional approaches in terms of utilization and served traffic. Moreover, to acquire the access network's requirement knowledge, we present a machine learning-based approach, which predicts the state of the network and is utilized to manage the available spectrum accordingly. Our comparative results show that, in terms of spectrum allocation accuracy and utilization efficiency, a well-designed machine learning-based bandwidth allocation mechanism not only outperforms common static approaches, but even achieves the performance (with a relative error close to 0.04) of an ideal dynamic system with perfect knowledge of future traffic requirements.

Contributions

[J2] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Dynamic spectrum allocation following machine learning-based traffic predictions in 5G" IEEE Access 9 (2021): 143458-143472. (Area: Telecommunications; Quartile: Q2 (36/91); IF: 3.367 (2020)). doi: 10.1109/ACCESS.2021.3122331.

[C3] R. I. Rony, Elena Lopez-Aguilera, and Eduard Garcia-Villegas. "Cooperative spectrum sharing in 5G access and backhaul networks." In 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 239-246. IEEE, 2018. doi: 10.1109/WiMOB.2018.8589187.

5.1 Cooperative spectrum sharing in 5G access and backhaul networks

To meet the expected QoS of 5G, several wired and wireless technologies are being considered as potential backhaul solutions for 5G. As discussed in Section 2.2., among wired solutions, optical fibre utilizing different access technologies, such as Gigabit PON, Ethernet PON (EPON), PtP, PtmP, Next Generation PON (NGPON), NGPON2, WDM-PON, provides the highest capacity and very low latency. However, wired options lack scalability and are costly for new deployments.

On the other hand, wireless backhaul options are less costly and their deployment is faster and easier (cf. Section 2.2). However, they are very vulnerable to environmental effects and often lack capacity. An attractive option for future wireless backhaul is mmWave, that operates in three different bands, 60GHz (V-band), 70/80GHz and 90GHz (E-band), offering data rates up to 10Gbps [12] [13]. Additionally, European Telecommunication Standards Institution (ETSI) recently started to work into D-band (141-174.8 GHz) for higher capacity. Advanced technologies, such as spatial multiplexing and beamforming can be used to improve the overall performance of mmWave.

From the discussions therein, it is clear that the transport network will be a dominant element of 5G networks, which needs to be cost and resource-efficient. Acknowledging this condition, cooperative access-backhaul mechanisms ensure the efficient use of precious resources, where both networks are dependent on each other's requirements and constraints. Moreover, cooperative access-backhaul mechanisms have the potential to minimize network CAPEX and OPEX [C1]. Therefore, in this section, we first provide a discussion on the ben-

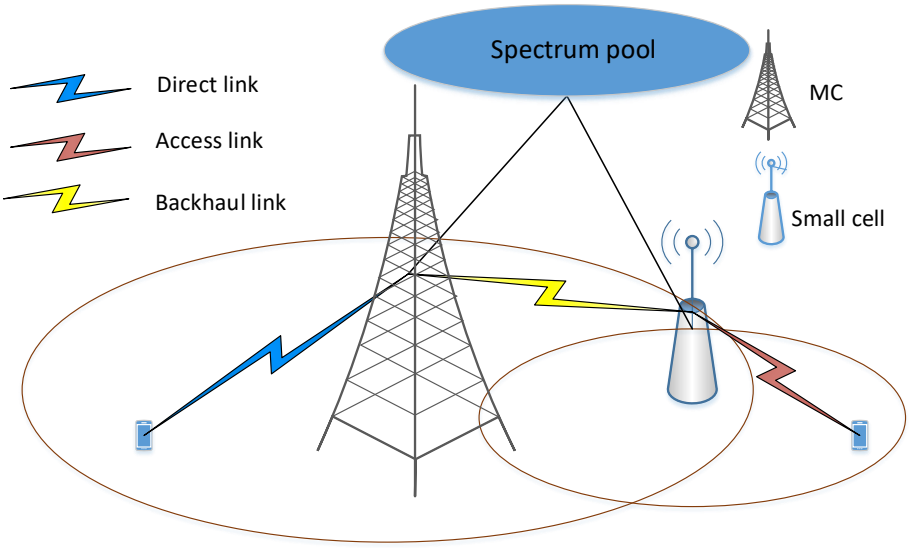

Figure 5.1: Three types of DL links under a common spectrum pool in a two-tier network.

efits of cooperative access-backhaul resource management and present cooperative spectrum sharing as one promising cooperative access-backhaul mechanism. Subsequently, we discuss and evaluate different spectrum sharing approaches through simulation.

5.1.1 Access-aware cooperative spectrum sharing

Discussion and the presented related work in Section 2.4.2 on *joint spectrum sharing* propose spectrum partitioning either in a two-tier network, i.e., among MCs and SCs, or between access and backhaul links of SCs. Considering the latter, few works suggest that resources to build the backhaul links can be partitioned either from the ones dedicated for SCs or from the corresponding to MCs. Focusing towards 5G dense networks, we consider three types of DL links: (i) ***Direct link***, the link connecting UE to MC, (ii) ***Access link***, the link connecting UE to SC, (iii) ***BH link***, the link connecting SC to anchoring MC. This scenario requires flexible partitioning from a common spectrum pool, as shown in Figure 5.1. Reference [83] also considers three types of link to share spectrum, but employing HD-capable SCs, i.e., SCs can either transmit data to its UEs or receive data from corresponding anchor MC in a given time-frequency resource. That approach does not allow the full utilization of resources; rather, FD operation is preferred, although it might require intelligent techniques to take care of SI in the network. Hence, we consider both MC and SC are FD-enabled. Subsequently, we evaluate different spectrum partitioning approaches among the aforementioned three different DL links, where five possible scenarios, depicted in Figure 5.2, are identified according to the

state-of-the-art. Note that, in the following, we use α as the proportion of the bandwidth dedicated for direct links, β for the BH links, and γ for the access links.

Spectrum Sharing Approach-1 (SSA-1): Each AP (MC and SCs) in the network can access the full bandwidth from the spectrum pool. MC treats the BH links in the same way as UE links [83], and allocates the bandwidth accordingly. Thus, total bandwidth allocated for MC is distributed among BH links and direct links.

Spectrum Sharing Approach-2 (SSA-2): In this approach, all the links, i.e., direct links, BH links and access links, operate in an out-of-band fashion. Thus, 50% (assuming a typical deployment, where α is equal to 0.5) of the total bandwidth is used for direct links, and 50% is for SC networks. As this is a fully out-of-band approach, the dedicated bandwidth for SC network is further equally shared among the access links and BH links without spectrum reuse. Thus, $\alpha = 0.5$, $\beta = 0.5/(NSC*2)$, where NSC is number of SCs and $\gamma = 0.5/(NSC*2)$. A great benefit of this approach is that there is no interference in the network.

Spectrum Sharing Approach-3 (SSA-3): We consider that 50% of the bandwidth is dedicated for direct links, and that each SC reuses 50% of the bandwidth dedicated to SCs. Thus, all SCs' access networks are in-band, hence interfering to each other. The same 50% of the bandwidth is also used in the BH links, yet managed by the MC and hence shared in a 0.5/NSC fashion. Thus, $\alpha = 0.5$, $\beta = 0.5/NSC$ and $\gamma = 0.5$. In SSA-3, spectrum re-use allows higher bandwidth in SC's access links than in SSA-2. On the other hand, in this approach, access links interfere to each-other, and additionally, self-interference also affects the BH links.

Spectrum Sharing Approach-4 (SSA-4): Here, 50% of the total bandwidth is dedicated to BH network, and the remaining 50% is re-used by all APs, (i.e. MC's direct links and SCs' access links). Hence, BH network is totally out-of-band and does not experience any interference. On the other hand, all access networks (i.e. direct links and access links) are in-band, and thus, interfere to each other. Using this approach, we have, $\alpha = 0.5$, $\beta = 0.5/NSC$ and $\gamma = 0.5$, being the values of α, β and γ the same as in SSA-3, but showing a different interference scenario.

Spectrum Sharing Approach-5 (SSA-5): Similar to SSA-3, 50% of the total bandwidth is dedicated for SCs' network and remaining 50% is dedicated for direct links. However, unlike SSA-3, in SSA-5 all the SCs are out-of-band. Hence, dedicated 50% of the spectrum is distributed among the SCs, and each SC re-uses the same spectrum in both BH link and access link. In this approach, access links of different SCs do not interfere to each other, but, for each SC, corresponding BH link interferes with access link. BH links also experience SI. Thus, the distribution is $\alpha = 0.5$, $\beta = 0.5/NSC$ and $\gamma = 0.5/NSC$.

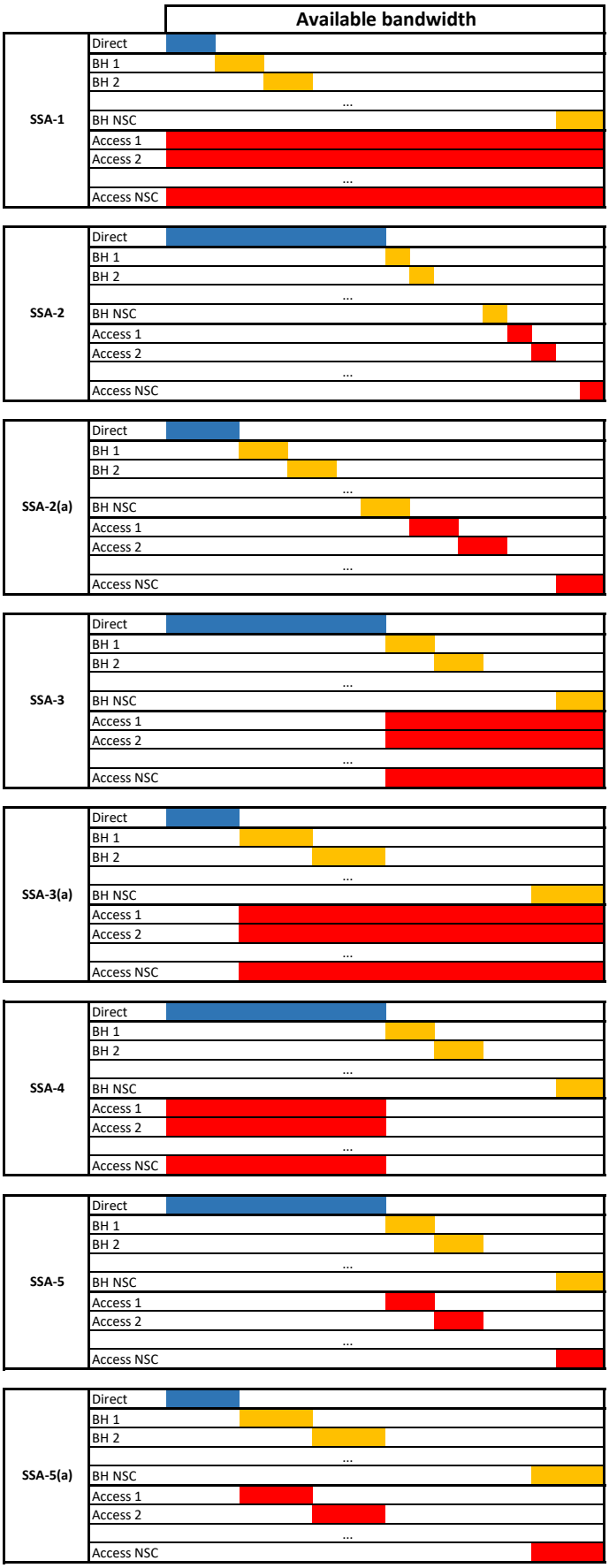

Figure 5.2: Set of Spectrum Sharing Approaches considered.

Table 5.1: Simulation Assumptions

Parameters	**Value**
Carrier Bandwidth (B)	100 MHz
Carrier frequency	3.5 GHz
Number of SCs	8 per MC
MC coverage radius	150m
SC coverage radius	15m
Minimum distance between MC and SC	20m
Minimum distance between adjacent SCs	15m
MC transmit power	49dBm in 20 MHz band
SC transmit power	30 dBm in 20 MHz band
Thermal noise level	-174 dBm/Hz
MC and SC noise figure	9 dB in 20 MHz band
Antenna gain	MC: 17 dBi; SC: 5dBi
Antenna height	MC: 25m; SC: 10m; UE; 1.5m
Channel model	3D model from [3]
Number of UEs	10 per MC; 5 per SC
SI cancellation factor (C_{SI})	100 dB [42]
Propagation type	BH link: LoS; Direct links: NLoS; Access links: LoS; Interferer links: NLoS
Number of simulations	1000 simulations, each with a random deployment of UEs and SCs

5.1.2 Evaluation and results

To evaluate different spectrum sharing approaches, we consider a dense urban scenario wherein a two-tier 5G network is deployed; one MC and several SCs cover the area (e.g., Figure 5.1). In such a network, a spectrum pool is managed and controlled by the central controller, which distributes the frequency resources among the potential links. The simulation assumptions summarized in Table 5.1 follow the use cases defined in [3] and the 5G deployment scenarios as predicted by METIS-II project [2].

As discussed throughout this thesis, resources are scarce, and thus, more efficient and fair distribution is required thereof. According to METIS-II recommendation for system level simulation, MC will serve 10 UEs, there will be 8 SCs per MC, and each SC will serve 5 UEs (cf. Table 5.1). Taking this into account, SCs will carry 80% of the UE traffic, whereas MC (direct link) will serve around 20%. With this in mind, we propose three additional SSAs.

Spectrum Sharing Approach-2(a) (SSA-2(a)): This is an access-aware version of SSA-2, where MC gets 20% of the bandwidth and SCs' network 80%, according to aforementioned traffic requirements. Thus, $\alpha = 0.2$, $\beta = 0.8/(NSC*2)$ and $\gamma = 0.8/(NSC*2)$.

Spectrum Sharing Approach-3(a) (SSA-3(a)): This approach consists in an access-aware version of SSA-3, following the same idea as for SSA-2(a). Thus, $\alpha = 0.2$, $\beta = 0.8/NSC$ and $\gamma = 0.8$.

Spectrum Sharing Approach-5(a) (SSA-5(a)): Similar to SSA-2(a) and SSA-3(a), SSA-5(a) provides an access-aware version of SSA-5. Hence, $\alpha = 0.2$, $\beta = 0.8/NSC$ and $\gamma = 0.8/NSC$.

The aforementioned eight SSAs are illustrated in Figure 5.2. We performed 1000 random simulations using the developed Matlab tool for the eight different approaches presented, and following the simulation assumptions exposed in Table 5.1. The achievable throughput for each link has been computed utilizing the Shannon - Hartley theorem (Eq. 5.1), where C is the channel capacity in bps, B corresponds to the channel bandwidth in Hz, and S is the received signal power in Watt. For in-band backhaul solutions, SI power (R_{SI}) in W has been considered in addition to the sum of co-channel interference power (I), and the sum of thermal noise and noise figure at the receiver (N) in W (Eq. 5.2). $\mathrm{R_{SI}}$ depends on C_{SI} value as $R_{SI} = P_{SC}/C_{SI}$ [59], where P_{SC} corresponds to SC transmitted power.

$$C = B * \log_2(1 + \frac{S}{N + I}) bits/s \tag{5.1}$$

$$C = B * \log_2(1 + \frac{S}{N + I + R_{\mathrm{SI}}}) bits/s \tag{5.2}$$

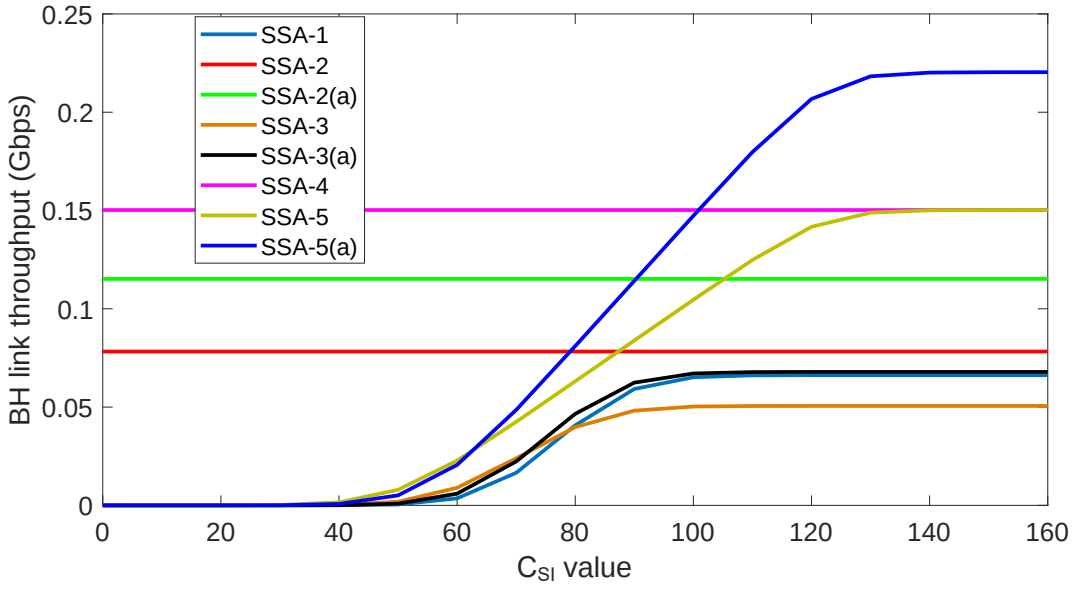

Figure 5.3: BH link throughput vs C_{SI} value.

Firstly, we show the impact of C_{SI} values on BH link throughput (Figure 5.3). As illustrated, SSA-2, SSA-2(a) and SSA-4 do not experience any dependency on C_{SI} value as these approaches use different band for BH and access links, thus not experiencing any SI.

On the other hand, for in-band approaches, the maximum achievable throughput improves with the increment of C_{SI} value. Obviously, with low C_{SI} values, in-band approaches perform worse than out-of-band solutions, even using a higher bandwidth share for BH links. Additionally, after certain C_{SI} values (between 80 and 120 dB), BH link achieve its maximum capacity. As suggested in [42], a value of $C_{SI} = 100$ dB is used hereinafter in those BH links experiencing SI.

Figure 5.4(A) provides average BH and access link capacity for the different SSAs. Maximum achievable capacity for SC's UE consists in the minimum value between access and BH link throughput. From the figure, it can be observed that, BH links are acting as the bottleneck in most of the cases (BH link throughput is lower than access link throughput). On the other hand, for SSA-5 and SSA-5(a), the BH link achieves higher average throughput in comparison to the access link, due to the high power received from the MC in BH interfering links (access links and BH links are in-band). Figure 5.4(B) illustrates maximum achievable throughput per UE for both MC and SC (note that, for SC UE, value corresponds to the minimum between BH and access link divided by the number of UEs per cell). Boxplot

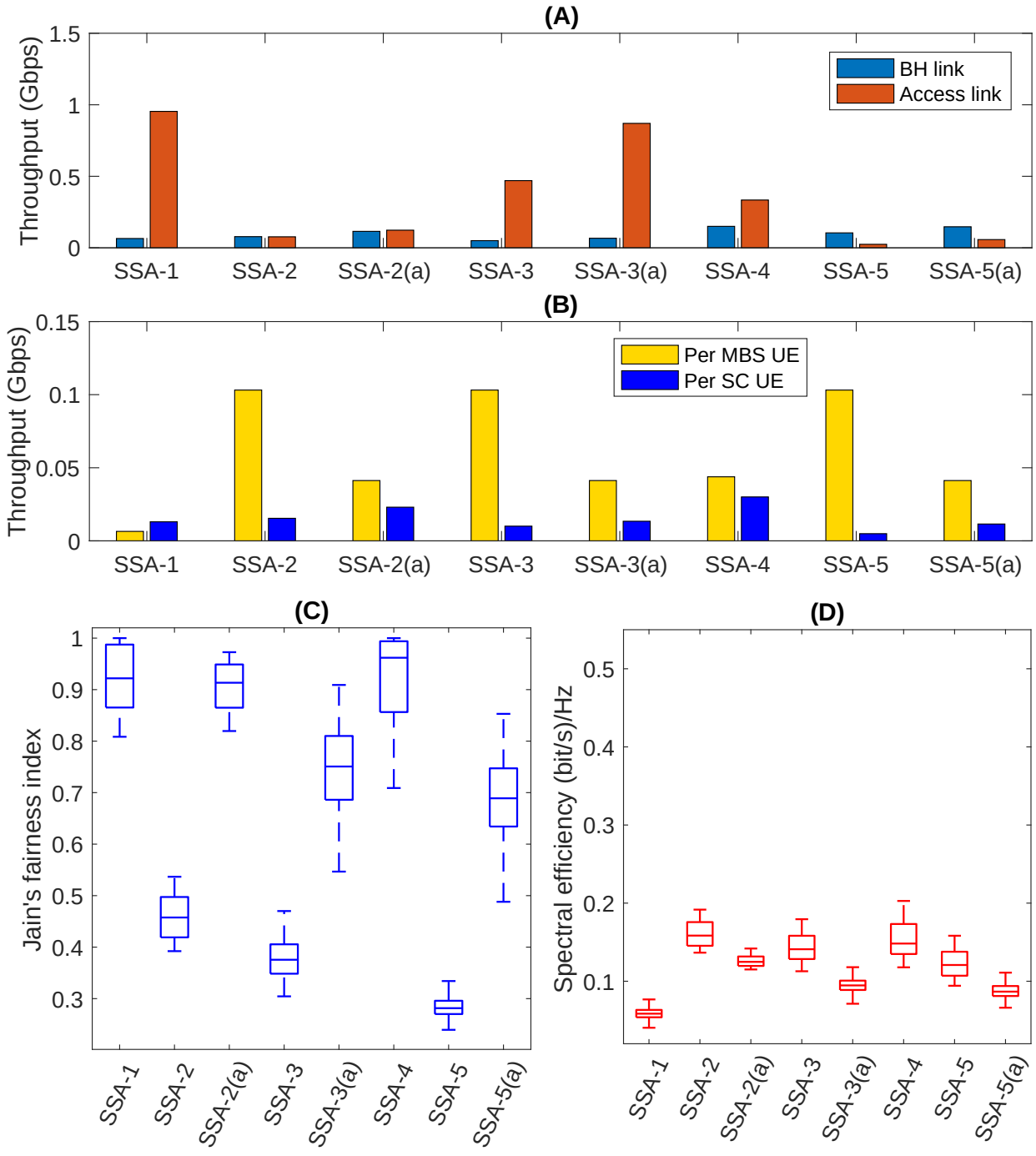

Figure 5.4: Performance evaluation of different SSAs: A) BH and access link achievable throughput; B) Per UE throughput of MC and SC; C) Jain's Fairness index; D) Spectral efficiency.

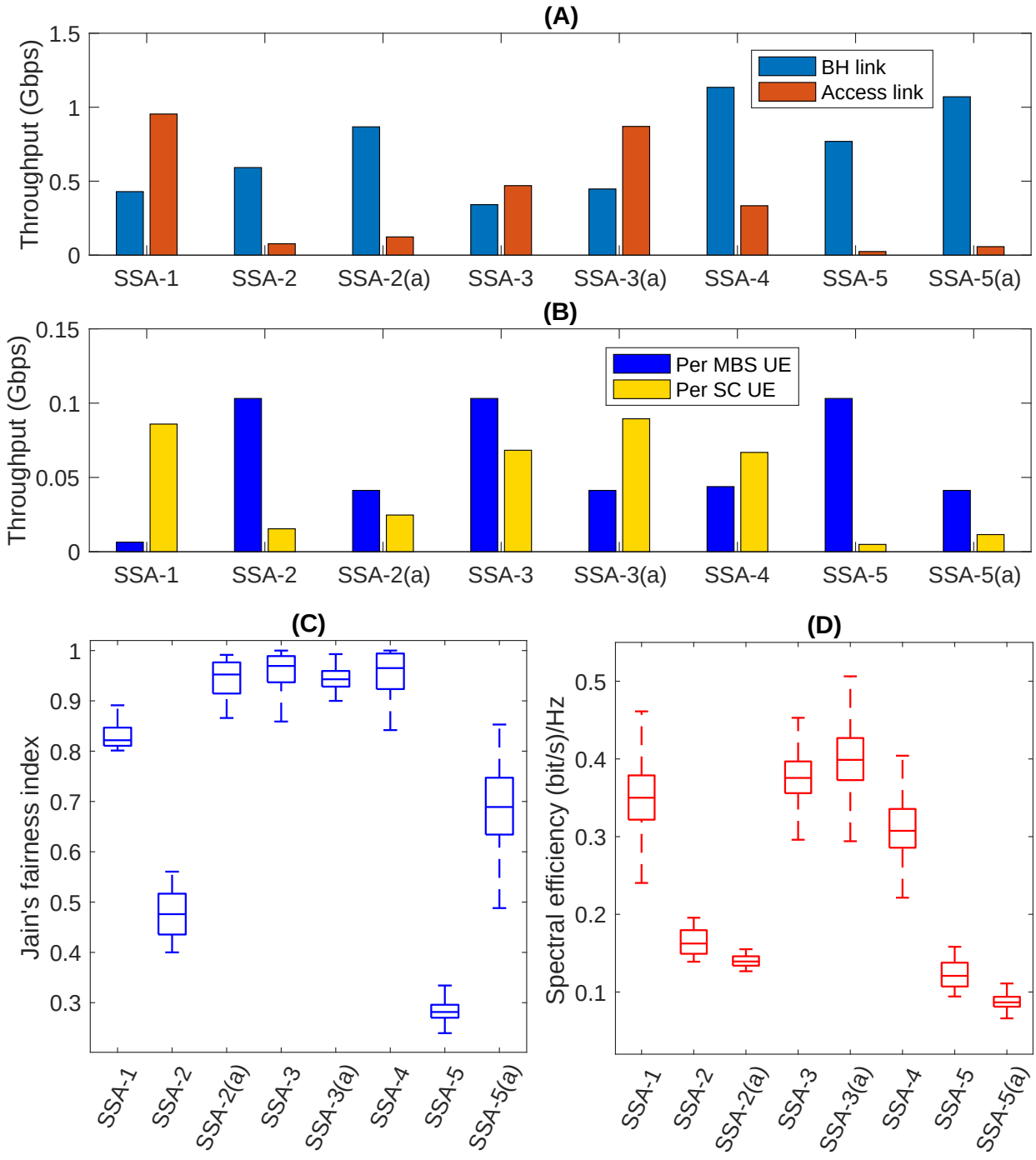

Figure 5.5: Performance evaluation of different SSAs with MIMO enabled in BH link: A) BH and access link achievable throughput; B) Per UE throughput of MC and SC; C) Jain's Fairness index; D) Spectral efficiency.

for 1000 simulations in Figure 5.4(C) represents *Jain's fairness index* calculated utilizing Eq. 5.3, where n is the number of UEs and X_i consist in the individual throughput of each UE [84]. Boxplot (1000 simulations) in Figure 5.4(D) depicts the system spectral efficiency (SE), which corresponds to the ratio between aggregated UE throughput in bps and total bandwidth in Hz. In the presented boxes, the lower bound and the upper bound of the boxes represent the 25th and 75th percentile of the data, respectively, and the line inside each box represents the median of 1000 simulations. The whiskers are extended to the maximum and minimum values inside the simulation tests.

$$F_\mathrm{j}(w) = \frac{(\sum_{n=i}^{n} X_\mathrm{i})^2}{n * \sum_{n=i}^{n} {X_\mathrm{i}}^2} \tag{5.3}$$

Evidently, proposed SSA-2(a), SSA-3(a) and SSA-5(a) show higher fairness than corresponding legacy approaches (SSA-2, SSA-3 and SSA-5, respectively). Although, in some cases, the whiskers show a larger window for the access-aware approaches, they always provide higher values, thus being fairer than their legacy counterparts. Allocating more resources into the SC networks, which suffer from a bottleneck in the BH link, allows more room for improvement, and thus, additional variability in the fairness enhancement is observed. On the other hand, SE in the fairest approaches is slightly lower. This is due to the fact that direct link gets less portion of bandwidth, whereas the larger bandwidth assigned to SCs is shared among eight BH links, and thus, the throughput increment in each SC's UE, does not compensate the loss of capacity experienced in direct link UEs.

As already discussed, BH links become the bottleneck in most of the cases and thus, maximum throughput for SC's UE is limited, although higher throughput is achievable in the access link. To tackle this problem, we study the effect of Multiple Input Multiple Output (MIMO) in the BH network. As recommended in [3] and [2], eight transmitter antennas and eight receiver antennas (8 x 8 MIMO) are considered to perform the MIMO operation. We use capacity computation from [85] (Eq. 5.4) to calculate the MIMO capacity in the BH links:

$$C = B * \log_2 \det(\mathbf{I} + (\frac{SINR}{n}) * \mathbf{H}.\mathbf{H}^{+}) bits/sec \tag{5.4}$$

where n is the number of transmitter/receiver antennas, $\mathbf{I}$ corresponds to the n x n identity matrix, $\mathbf{H}$[1] is the normalized channel matrix, which is frequency independent over the signal

[1]For the sake of simplicity, we build $\mathbf{H}$ is as an n x n random matrix following a complex normal distribution.

bandwidth, and $\mathbf{H}^{+}$ is the transpose conjugate of $\mathbf{H}$.

Figure 5.5 (A, B, C, D) represents the corresponding results with MIMO enabled in BH links, which shows that MIMO technology can be efficiently used to overcome the bottleneck situation in BH links. Unlike Figure 5.4(A), Figure 5.5(A) shows that, with MIMO enabled, the bottleneck has been moved towards the access link in most of the cases. As mentioned earlier, for approaches SSA-5 and SSA-5(a), BH is not the bottleneck, even without MIMO enabled (Figure 5.4(A)) and hence, enabling MIMO in the BH link does not have any impact on the presented results for these two approaches. This is due to the fact that, for these approaches, access link of SC limits the achievable data rate, and thus, requires MIMO operation in the access link to improve performance.

Note that when the capacity of the BH links is larger than the capacity of the access links, part of the resources granted to the BH will be wasted. Therefore, it makes sense to share BH resources with access links to increase efficiency (cf. SSA-3 and SSA-3(a)). When BH is the bottleneck, the highest efficiency is obtained when BH is isolated (SSA-4). However, when BH is not the bottleneck (e.g., because of MIMO), reuse in the SCs (i.e., SSA-3(a)) outperforms other approaches.

With MIMO enabled in the BH links, we present preliminary results, which can be further exploited considering other potential benefits of MIMO, such as interference mitigation. Additionally, MIMO can be also useful in the access link in order to increase corresponding capacity. However, if the perfect balance between BH and access links is not achieved, resources can be wasted in either link, as throughput experienced by SC's UE is always limited by the minimum value between BH and access link capacity. And thus, to achieve the perfect balance, intelligent cooperative optimization of access and BH is essential.

Cooperative access-backhaul mechanism is a key enabler of 5G, which can make the future networks more feasible, flexible, resource and cost-efficient. Discussed results put additional weight to the aforementioned statement. Moreover, cooperative access-backhaul design and optimization allow to relax BH requirements and validate the idea of offering an on-demand BH service.

There are some additional lessons learned from the presented results. Firstly, the value of C_{SI} has a great impact on the performances of in-band solutions. Thus, to benefit from the idea of in-band allocation of spectrum, self-interference has to be taken carefully into account. Secondly, access-aware spectrum allocation of SCs and MBSs provides a fair distribution of resources among different UEs at the cost, in some cases, of reducing spectral efficiency. Finally, depending on the location of the bottleneck (access or BH), different spectrum sharing strategies should be chosen.

5.2 Dynamic spectrum allocation following machine learning-based traffic predictions in 5G

In this section, we focus on spectrum allocation techniques to base stations in future 5G networks according to their access network dynamic requirements. To ensure the best utilization of the available spectrum, we propose an allocation mechanism that follows the instantaneous capacity requirements in the access network. Additionally, we analyse the adaptation of a ML technique to predict the capacity requirements throughout different times of the day. First, we review the state of the art by presenting a detailed discussion on different spectrum allocation techniques, and the growing popularity of different ML techniques in wireless networks. Subsequently, we describe the proposed intelligent spectrum allocation technique. Afterwards, the description of the evaluation scenario, the related simulation assumptions and comparative traditional approaches are provided. Finally, detailed analysis and comparative results for the proposed intelligent access-aware spectrum allocation technique are presented. Additionally, we discuss the performance analysis of the ML-approach, which provides bandwidth assignments according to traffic predictions.

5.2.1 Spectrum allocation techniques

Licensed spectrum has always been the most expensive and scarce resource from wireless networks service providers' point of view. With the aggressive growth of capacity demand on wireless networks, the spectrum became more precious, given that the most straightforward way to meet the higher capacity requirements is to increase the assigned BW [86]. However, this approach also brings additional challenges. BW distribution of the spectrum among different cells for ensuring the best utilization of scarce resources has always been a very challenging task.

Throughout the previous generations of cellular networks, Fixed Spectrum Allocation (FSA) has been a popular approach for BW distribution among different APs. In this solution, an initial capacity plan is performed, and BW is allocated to the APs according to the maximum requirements it is expected to serve during peak hours, and to the available resources. Later on, BW allocation remains static and agnostic to the dynamic capacity requirements of the different APs in the network [87].

Compared to FSA, in Dynamic Spectrum Allocation (DSA), allocation of spectrum follows the instantaneous requirements of the APs, according to the available resources. A simple approach is presented in [87], where a DSA is triggered after a particular time interval, which estimates the load for the next time interval, computes the spectrum requirements

and, if needed, allocates additional spectrum for the subsequent time interval. In [87], authors argued that DSA can be potentially used to share the idle spectrum in a multi-operator environment sharing a spectrum pool.

DSA technique has been very popular since its inception as it increases the spectral efficiency of the network by allowing the under-utilized frequency bands to be employed in an efficient way [88]. With the evolution of wireless technologies, DSA techniques also evolved. For example, different approaches emerged in the Cognitive Radio (CR) environment, where DSA is used to allocate the idle channels to Secondary Users (SU), which usually are unlicensed and have lower priority in the network [89–93], based on Fuzzy logic, Q-learning, randomized rounding algorithm, etc., to learn, estimate and allocate the required spectrum.

In [94], authors proposed a Reinforcement Learning (RL)-based DSA technique for spectrum allocation to the IoT users in a cellular network. In this proposal authors successfully showed that the DSA technique can be used to identify the underutilized spectrum in the network to be reused for a sensor-aided IoT network, subsequently, enhancing the spectrum re-usability.

In [95], RL-based DSA technique is used, where each cell can take its own decision of spectrum allocation to its users with the objective of maximization the overall SINR. In this work, the authors tested the DSA technique in a decentralized approach, where each cell acts as an individual DSA agent, although collecting the spectrum allocation information from the neighbouring cell or DSA agent. In another work [96] from the same authors, RL-based DSA technique was tested in a centralized approach, where a central DSA agent controls and takes the spectrum allocation decision for all the cells in a considered area. The two studies showed that DSA-based spectrum allocation technique can outperform traditional FR techniques, both in a homogeneous (i.e., macrocell only), or heterogeneous (coexistence of macrocell and femtocell/SC together) scenario. However, both the presented works utilize DSA technique to minimize intercell interference in OFDMA networks. In this section, however, we propose a technique similar to DSA, where spectrum allocation to the cells follows the access network's dynamic requirements of each cell in a real network scenario.

ML-based Dynamic Frequency and Bandwidth Assignment (DFBA) focusing on spectrum allocation to the small cells in a cellular network was studied in [97]. The authors presented a technique to learn and predict LTE KPIs (e.g., SINR per resource block, MAC level throughput, delay, etc.) and assign/rearrange spectrum allocation to the LTE-based SCs in the network. However, this work also does not consider the opportunity to enhance spectral efficiency and system fairness by allocating the scarce spectrum according to the current load of the cells.

During the previous generations of cellular networks (i.e., 1G and 2G), compilations of a

set of neighbouring cells were known as clusters, and the entire available BW was distributed among the cells within a cluster, avoiding overlapping portions of the spectrum. Thus, the whole available spectrum could be reused again in each cluster [98]. On the other hand, in newer generations of wireless networks, i.e., 4G and 5G, as part of the strategy to reach the required capacity, the technology allows the use a FR factor equal to one [98], which means that the same frequency band is reused in each cell in the network. The interference introduced due to such aggressive reuse is expected to be handled (i.e., controlled down to a minimum level) by new generation of advanced technologies, such as, eICIC, beamforming, etc. However, reusing the entire BW is neither efficient nor a fair approach since, in dynamic 5G networks, different APs will have distinct levels of load to serve.

Moreover, in future UDN, due to the wide variety of use cases, different sizes of APs (e.g., MBS, SCs) will serve different numbers and types of users. Additionally, due to the mobility and higher user density, the capacity requirements of each AP can differ largely. Moreover, the number of served users or the amount of traffic carried by each AP will vary dynamically with the time of the day [99]. Thus, in 5G networks, the static allocation of BW is not an efficient approach, rather a dynamic solution is required, where, unlike the discussed related work, BW distribution must follow the current requirements of individual APs. In this work, utilizing a spectrum pool, we adopt the DSA technique to allocate required BW to MBSs according to their current requirements, which, in our study, are based on real traffic traces. A similar approach was presented in [100], where authors address the problem of dynamic changes in the required capacity in a multi-service (i.e., cellular network, vehicular network and IoT) 5G network. In [100], authors presented how underused frequency bands from one service, e.g., IoT networks, can be requested and used by another service, e.g., vehicular networks, to meet individual dynamic capacity requirements during a congestion situation. Another approach of spectrum sharing can be found among coexisting different Mobile Network Operators (MNOs) to enhance the capacity of their network [101], [102].

In this work, however, we consider each cell belonging to a single MNO, as an independent entity having its own dynamic capacity requirements. A spectrum pool, which we consider is managed by an intelligent SDN-based controller (similar to the centralized DSA approach in [96]), allocates and rearranges small portions of licensed spectrum bands (e.g., 5 MHz chunks) extracted from the entire available BW (e.g., 100 MHz). Even though 100% frequency reuse is allowed (i.e., FR=1), our approach seeks to meet each cell's requirement while minimizing the bandwidth assigned. In this way, we also mitigate the effect of that aggressive frequency reuse, minimizing interference without sacrificing capacity. Additionally, utilizing a ML technique, we predict the throughput requirements of each cell and allocate BW according to the predicted behaviour. To validate the approach, we also compare real and predicted

throughput requirements, and compare the BW allocation for real and predicted approaches. Such an approach can be further studied in a multi-MNOs scenario, which we leave for future work.

5.2.2 Machine learning in wireless networks

The benefits of using ML techniques in wireless communications can be multi-fold. Utilizing ML techniques, a network is able to analyze the behaviour, learn from it, predict future status, and prepare itself for it. In wireless networks, by relating the system parameters to the desired objective [103], ML techniques can address the challenge of traditional optimization approaches, which tend to leave a large gap between the theoretical and real-time design of the network [104]. Therefore, for the development and automation of 5G networks, ML-based approaches are getting enormous attention and can be adopted in different aspects of cellular networks, e.g., interference management [104], beamforming [105–107], link quality estimation [108], 5G-based IoT [109], energy efficiency [81, 110], resource management [111], etc. There are different ML techniques: i) supervised learning, where the model learns by studying a set of labeled data, ii) unsupervised learning, where the model learns from a set of unlabeled data and, iii) reinforcement learning, where the model learns by assigning positive and negative rewards for its actions. Their corresponding different learning models (e.g., Support Vector Machines (SVM), K-means clustering, Gradient follower (GF), etc.), and their usage guideline for future wireless networks are well summarized in [112–114].

In this study, unlike the DSA techniques discussed earlier, we decouple the ML and DSA agents. We use a ML technique to predict the traffic, i.e., access network requirements, since reliable traffic prediction is already being considered as a key enabler for future wireless networks to improve QoS by reducing uncertainty [115]. The central controller collects the predicted traffic requirements of the access network for each cell from ML agent and, after necessary calculations, it (the central controller) assigns/rearranges the spectrum allocation to each cell accordingly. In this way, the central controller offloads the ML complexity into a dedicated separated agent. Additionally, as discussed, most of the DSA techniques use RL-based learning, which requires a relatively large time, since it is a positive/negative reward-based process. Moreover, unlike RL-based DSA agents, our proposed central controller takes a faster decision following predicted requirements, since the proposed technique restricts a large number of spectrum allocation combinations among the cells by enforcing a few realistic constraints, as discussed in the next section. Therefore, our proposed spectrum allocation technique requires a simple time-series predictive ML tool.

In [116], utilizing collected trace data, authors have developed a model to predict the ag-

gregated traffic and, subsequently, reduce the monitoring effort. Classification and clustering of cells were used to improve the traffic predictions in [81, 117]. However, these approaches require additional cell-level data. On the other hand, advanced ML-techniques such as Neural Networks (NN) reduce the dependency on additional features, and can be efficiently used to predict the traffic by analyzing a time series traffic data. A NN (aka. Artificial Neural Networks (ANN)), is a supervised ML technique, which emulates the way human brain works, by simulating artificial neurons with basic functionalities created from the complex computation [113]. A typical NN is composed of three types of functional layers, i.e., an input layer, one or more hidden layers, and an output layer. All the layers consist of a set of nodes, which are connected with adjustable weight coefficients to each of the nodes in the next layer. The weights connect the input data (via input layer) to the activation and transfer function (inside hidden layer) and generate an output (in the output layer) if the weighted sum activates the neuron [118]. Since NN can solve non-linear complex problems by finding true relations between the input and output parameters, and confirms the maximum level of generalization, NNs are very commonly used as a ML technique in recent studies [119–127].

In [128] authors have summarized a few studies on the benefit of NN network-based spectrum prediction techniques, and concluded that without needing much prior knowledge of the system, ANN shows the best performance.

Long-Short Term Memory (LSTM), a kind of recurrent ANN, avoids long-term dependency on input [129] by using additional information about whether to remember or forget it. In [129], authors have presented LSTM-based multi-step traffic prediction of a LTE network. However, LSTM is more complex to implement compared to the classical NN. On the other hand, LSTM is more suitable to go beyond the available time series data and predict traffic requirements for one or more time-steps in the future. In this study, however, we focus on traffic prediction during the available time series, which can be validated on the existing data. According to [130], Autoregressive Neural Network (NARNET), an NN-based ML technique, is arguably one of the best ML techniques to predict non-linear time series data. Additionally, in [131], authors showed that NARNET outperforms other studied techniques in terms of Coefficient of determination or R-squared (R2), which implies higher reliability for forecasts. More detail about NARNET is discussed in Section 5.2.4.

5.2.3 Access-aware dynamic spectrum allocation

As discussed in Section 5.2.1., following a static approach, spectrum is allocated to the APs without timely knowledge of the varying access network requirements, and it is usually based on the expected traffic during the peak hour. Subsequently, in most of the cases,

either the spectrum is over-provisioned or under-provisioned [99]. In a dense 5G network, due to the aggressive capacity demand by a large number of connected devices, such a static approach for spectrum allocation becomes unfair. Therefore, access-aware spectrum allocation, which allocates the spectrum to each of the APs following current requirements, is unavoidable. Our work presented in Section 5.1 discusses such an awareness approach, where access-aware spectrum allocation is applied to a two-tier heterogeneous network (i.e., the coexistence of MBS and SCs), and SCs are anchored to the MBSs via wireless links. In such a complex network, three different types of links compete for the same pool of spectrum resources: i) direct link, the link between MBS and its users; ii) backhaul link, connecting the SCs to the anchored MBS and, iii) access link, the link between SC and its users. In [132], authors state that, in comparison with the traditional approaches (i.e., unaware of access demands), the aforementioned access-aware spectrum allocation solution ensures a better network performance in terms of user throughput, spectrum allocation fairness and spectral efficiency. However, the access-aware approach presented in [132] can still be considered static, since the dynamic changes in the access network requirements are not taken into consideration. In this work, we propose a dynamic access-aware spectrum allocation, where BW allocation is performed following the dynamic changes in the access networks' requirements varying with time. Unlike [132], this work is based on a real deployment (cf. Section 5.2.4.), where all APs can be considered, in fact, as MBS. Each MBS in this realistic model constitutes (one or) multiple cells. For this reason, the term MBS is henceforth used instead of the more generic AP.

5.2.3.1 System model

We assume, BS is the set of MBSs, each having nMC_i cells, so that $\sum_{i=1}^{numMBS} \boldsymbol{nMC}_{\mathrm{i}} = \boldsymbol{numMC}$, where $numMBS$ is the number of MBSs, and $numMC$ is the total number of cells. MC_i is then the set of nMC_i cells of the i^{th} MBS. B_i is the total BW assigned to that MBS, so that $B_i = \sum_{j=1}^{nMC_{\mathrm{i}}} b_{\mathrm{ij}}$, where b_{ij} are the elements of matrix B, representing the BW assigned to j^{th} cell in i^{th} MBS. The minimum frequency resource (i.e., 5 MHz in our scenario) is defined as b_m, and b_M is the maximum BW (i.e., 100 MHz in our scenario). Finally, R is an array, the elements r_{ij} therein represent the capacity requirements for j^{th} cell in i^{th} MBS, as read from the traces, or predicted by the ML model (c.f. Section 5.2.5). Similarly, c_{ij} are the elements of the matrix C, representing the capacity available in j^{th} cell of i^{th} MBS, according to the currently assigned bandwidth, and following Eq 5.1.

For a given amount of assigned BW, to compute the achievable throughput (C) in bits per second (bps) for a cell, we utilize the Shannon - Hartley theorem (Eq. 5.1). As explained

in Section 5.1.2, B is the total allocated BW for a cell (b_{ij}) in Hz, S is the received signal power from the serving cell, which is calculated subtracting the pathloss from the transmitted power and antenna gains, I is the received power from the interfering links, and N is the sum of thermal noise and noise figure. Note that the received signal and the interference are obtained after 1,000 random droppings of one UE within the cell's coverage area. Thus, both S and I refer to average values throughout the cell's coverage area, ensuring the generality of the results.

5.2.3.2 Proposed algorithm

Our proposed access-aware BW allocation algorithm is triggered at time intervals. It starts by sorting, in descending order, all the cells according to their requirements, so that the cell serving higher traffic has more chances to be assigned more resources in a BW-limited scenario. Afterwards, for each MBS, and for each cell in that MBS in the BS set, two conditions are checked; i) if the achievable capacity c_{ij} is less than the required capacity r_{ij} and, ii) if the anchoring MBS of that cell still has frequency resources available. Here, to limit the co-channel interference with the adjacent cell, which degrades the network performance due to strong coverage overlapping for the same frequency band, we avoid frequency re-use within a MBS but it allows the assignment of frequency resources up to the system bandwidth. Thus, the cumulative allocated BW of all cells belonging to one MBS (B_i) can not exceed the system BW (i.e., b_M). It can be reused in the other MBS. If both the aforementioned conditions for a cell are met, the cell is assigned with an additional minimum size frequency resource ($b_m = 5$ MHz, for example) taken from the entire available system BW (e.g., $b_M = 100$ MHz), as explained in Section 5.2.1. Hereafter, allocated BW to that cell and to the anchoring MBS is updated and the achievable capacity c_{ij} is re-calculated. If the new achievable capacity reaches the required capacity r_{ij}, the cell is considered as satisfied and counted as allocated for this time interval. On the other hand, if the new achievable capacity still does not meet the required capacity (let us call it unsatisfied cell), the cell waits for the next iteration, where according to the aforementioned two conditions, only the unsatisfied cells of the MBSs with available BW are assigned additional frequency resources. On the contrary, if the cumulative assigned BW for the MBS, which is anchoring the unsatisfied cell, reaches the maximum allowed BW, the cell is counted as allocated, whether it is satisfied or not. The iteration stops for this time interval when all the cells are considered as allocated.

The discussed BW allocation approach is depicted in Algorithm 2.

Algorithm 5.2 Access-aware dynamic BW allocation

```
1: Input: numMBS, numMC, nMC, B, R
2: Initialize B_i ← 0 for all i in (1, numMBS)
3: Initialize b_ij ← 0 for all i, j
4: allocated ← 0
5: while allocated < numMC do
6:    for all i in (1, numMBS) do
7:       Sort MC_ij according to their requirements r_ij in decreasing order
8:       satisfied ← 0
9:       for all j in (1, nMC_i) do
10:         if B_i + b_m ≤ b_M and r_ij > c_ij then
11:            b_ij ← b_ij + b_m; allocate additional b_m to cell MC_ij
12:            B_i ← B_i + b_m; update total assigned bandwidth to MBS_i
13:            c_ij ← capacity(MC_ij, b_ij);    update MC_ij's capacity according
   to Eq. 5.1
14:            if c_ij ≥ r_ij then
15:               allocated++
16:               satisfied++
17:      if B_i == b_M then
18:         allocated ← allocated + nMC_i - satisfied
```

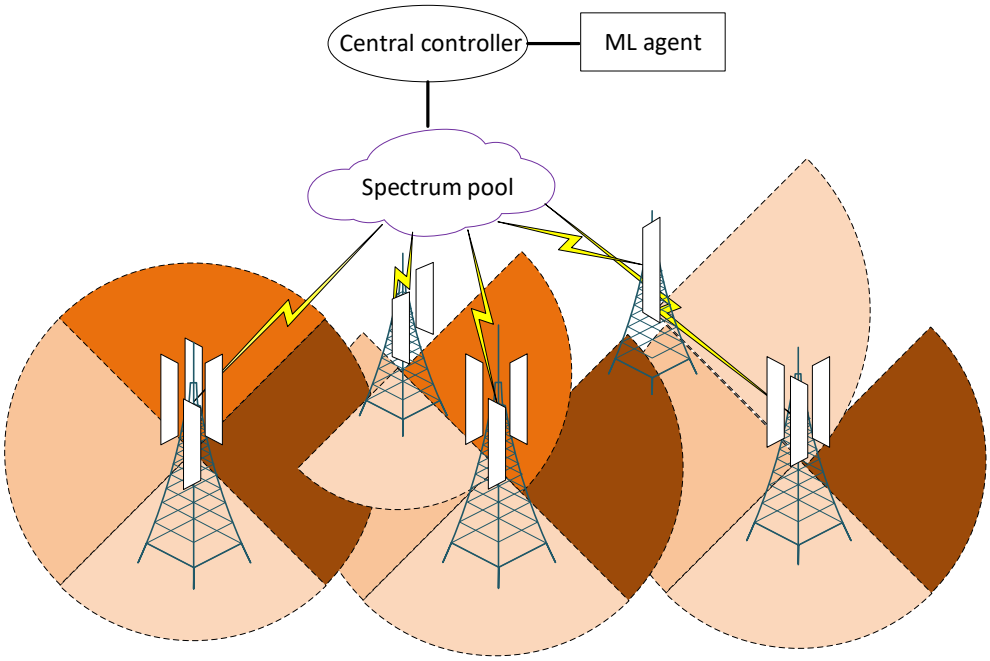

Figure 5.6: Spectrum pool concept to share the available system bandwidth.

5.2.4 Evaluation scenario

At this point, we want to evaluate the proposed access-aware dynamic BW allocation approach following a real traffic scenario. We use network traces provided by a real network operator, including data collected during seven days with 15 minutes granularity of a LTE based network serving a large number of users in the downtown area of a Greek city. The network consists of 21 eNBs, comprising 96 cells, with an average of 4.5 cells per eNB (varying from 1 to 9 cells). A simple representation of the considered scenario is depicted in Figure 5.6.

To observe the behavior of the traffic, we take one cell to start our analysis. Figure 5.7 depicts the average (i.e., average of seven days converted into a 24 hour scenario) Downlink (DL) data (in Mega Bytes (MB)) for a randomly selected cell varying with the time of the day. It can be observed from Figure 5.7, that the demand on DL data in the network varies largely with time and thus, a static BW allocation approach designed to meet those varying requirements is not an efficient approach, rather, a dynamic approach is required.

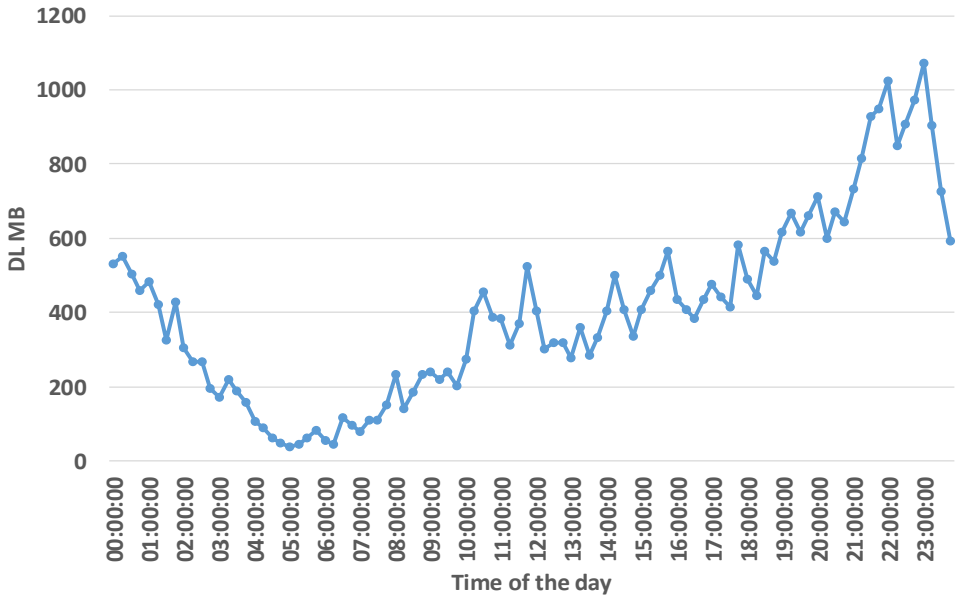

Figure 5.7: Time vs. average DL (in MB) of a randomly selected cell.

5.2.4.1 Simulation assumptions

To evaluate our proposed access-aware dynamic BW allocation approach using a real traffic scenario, we reproduce the network scenario, where the traces were collected, into Matlab code. Since our vision is to test our proposal in a 5G-like scenario, we make several assumptions, as recommended by the 5GPPP in [2], yet, assuming the traffic pattern follows the real LTE traces collected.

In [2], 5GPPP claims that in 5G networks, the most popular RAT for MBS will be Sub-6 GHz: Carrier frequency (CF) at 3.5 GHz with 100 MHz channel BW. Additionally, 3GPP and ETSI also identify Sub-6 GHz (CF: 3.5 GHz, BW: 100 MHz) as a candidate for 5G New Radio (NR) in [133]. Utilizing Sub-6 GHz operating band, we consider the narrowest transmission BW for a MBS to be 5 MHz, as recommended in [133], which we allocate in every iteration of the proposed BW allocation technique (cf. Section 5.2.3).

As mentioned earlier in Section 5.2.3, we consider a deployment with no frequency reuse within each MBS; that is, bandwidth assigned to the different cells inside the same MBS does not overlap. Additionally, we consider that directivity of antennas in the different cells and frequency assignment [2] are such that cells from adjacent MBSs do not interfere each

[2]Note that, our mechanism allocates bandwidth, but frequency assignment is out of the scope of this proposal and left for a future work

other significantly, and only one cell per non-serving non-adjacent MBS is considered as a source of interference for the application of Eq. 5.1. Thus, each cell will be experiencing interference from 17 other cells.

In the following, we describe three different BW allocation approaches to present our evaluation results in the subsequent section.

- *Static approach 1*: Each cell in a MBS is allocated a non-overlapping 20 MHz channel (the maximum channel BW for LTE [134]), while not exceeding the maximum BW per MBS, which is 100 MHz. If, due to the higher number of cells, 20 MHz per cell is not achievable, the entire available BW is evenly distributed among all the cells in that MBS, which is also a very common approach, where FR = number of cells (usually 3) [96]. For example: If MBS-1 has two cells, each cell will be allocated 20 MHz separated channels, and the remaining 60 MHz will not be assigned in MBS-1. On the other hand, if MBS-2 has 6 cells, each cell will be allocated 16.6 MHz.

- *Static approach 2*: The available BW is equally split among the cells, not limiting the maximum BW per cell to 20 MHz. In this way, each cell from MBS-1 in the previous example will be allocated 50 MHz. The allocated BW for the cells in MBS-2 remains the same as in static approach 1.

- *Proposed dynamic approach*: It follows the operation explained in Section 5.2.3; that is, iteratively, we allocate additional 5 MHz to every cell until its requirements for the next interval are met, or the maximum BW (i.e., 100 MHz per MBS) is reached.

5.2.4.2 Preparation of the Neural Network

According to the discussion presented in Section 5.2.2, given our data set and the objective of the prediction, the most suitable architecture is defined by NARNETs. We focus on a BW distribution problem that must follow the dynamic requirements of a cell, which changes with the time of the day. Hence, we follow a time series data, which are collected from a real network, to understand the dynamic changes in the required throughput.

In both closed loop and open loop scenarios, utilizing past samples (e.g., *y (t-1), y (t-2)..y (t-d)* of the output, known as delays (d), and by modeling the underlying characteristics of the time series, NARNET has the ability to self-learn and provide good multi-step predictions ($y'(t)$) of a non-linear time series [135–139]. A simple representation of NARNET model is presented in Eq. 5.5, where function f is the result of the process represented in Figure 5.8. Each neuron performs a linear combination of all its inputs, applying an adjustable weight

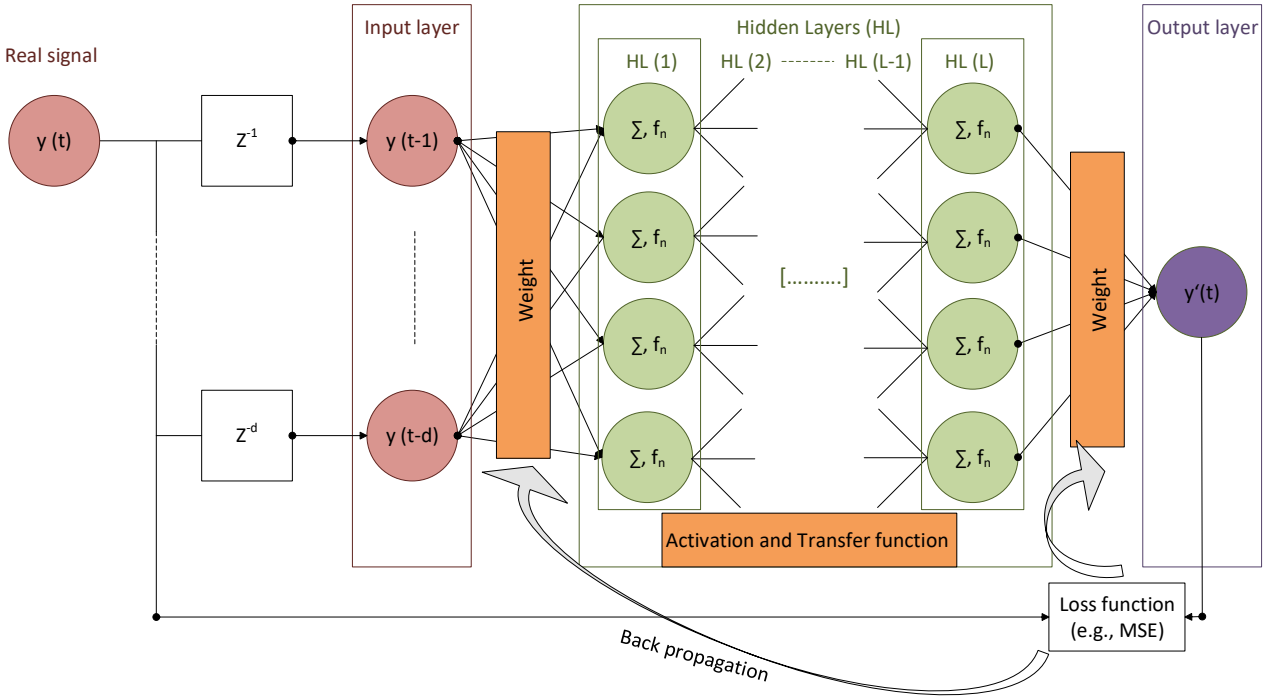

Figure 5.8: Structure of a typical NARNET.

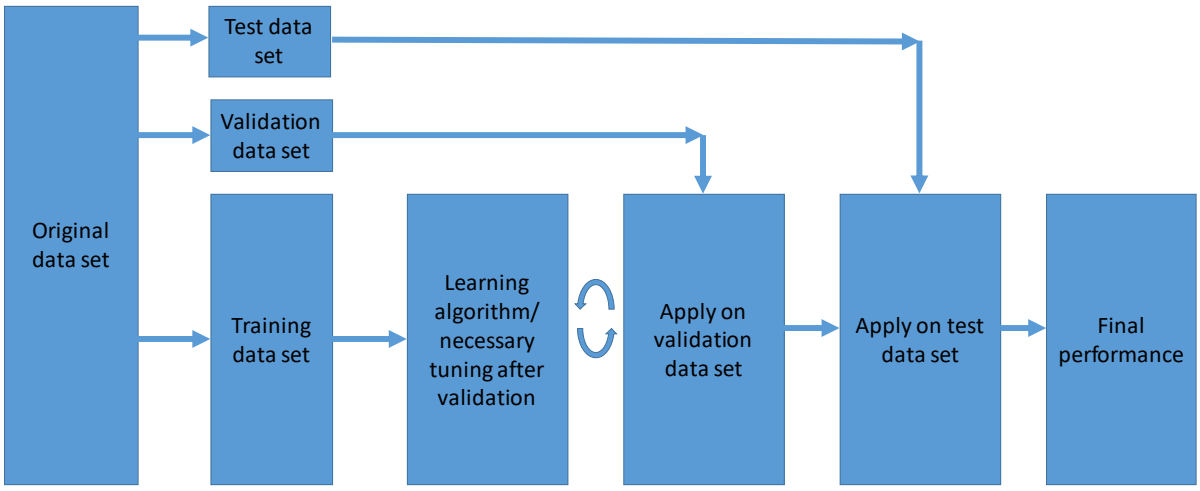

Figure 5.9: Basic representation of training, validation and testing process of NARNET

to each one, followed by an activation function (i.e., non-linear transformation) f_n.

$$y'(t) = f(y(t-1), y(t-2).....y(t-d)) \tag{5.5}$$

To predict the throughput requirements of the access network for each cell, we used NARNET. More precisely, we use DL MB (cf. Figure 5.7) data extracted from the traces as the input *y(t)* for the NARNET. To predict and compare the results with real traces we used Matlab. While creating a NARNET in Matlab, it allows the selection of the portion of the data set to be utilized during the three phases of learning, i.e., training, validation and testing. During the training phase, a part of the available data is used, and the weight coefficients are adjusted by taking feedback, known as backward propagation, from the comparison between predicted and expected output by means of a given loss function, e.g., Mean Squared Error (MSE) or other estimators. During the validation phase, the training is validated on a different data set, and the results are utilized to generalize (i.e., performance for unseen data) or improve the generalization of the NN. Finally, during the testing phase, the model is tested on the last portion of the data, and the NN performance is computed. A basic representation of the aforementioned process is depicted in Figure 5.9. Figure 5.8 presents the structure of a typical NARNET. Here, after trying different combinations, we end up using 70% of the data for training, 15% for validation, and 15% for testing, since that was the partition providing the best performance (i.e., least MSE) for our data set. Three different training algorithms are offered by a Matlab-based NARNET: i) Levenberg-Marquardt (LM), which was originally designed for faster results, however, consumes more memory and is less accurate; ii) Bayesian Regularization (BR), which has the additional objective of minimizing the sum of squared weights and, subsequently, achieves a good generalization of the model [140]. Compared to LM, BR is a slower model, however, it provides better generalisation for noisy data; and iii) Scaled Conjugate Gradient (SCG), which stops the training when generalization does not improve anymore, consumes less memory, and follows minimization of MSE as the only objective function. A detailed comparative analysis of LM and BR was presented in [140] and the results show that BR outperforms LM for different types of data sets. To find the best configuration for our data set, we performed NARNET training utilizing multiple combinations of the following hyper parameters: number of hidden layers (from 1 to 60), delay samples ($d = 1$ to 5) and training models (i.e., LM, BR and SCG). Following different combinations of the aforementioned parameters, we had a large number of predicted data sets, and we measured their corresponding performance by calculating Relative Error (RE) (Eq. 5.6) for each sample, along with MSE measured by NARNET. Finally, the best combination of the different hyper parameters in NARNET for our data set consists of 8 hidden layers, delay of 1 sample and BR as the training algorithm,

Table 5.2: Simulation assumptions following 5GPPP recommendations [2] [3]

Parameters	**Value**
Number of MBSs	21
Number of Cells	96
RAT	Sub-6 GHz
Carrier Bandwidth (BW)	100 MHz
Carrier frequency	3.5 GHz
BW of a single frequency band	5 MHz
MBS coverage radius	200 - 800m (Depending on the location of the MBS)
MBS transmit power	49dBm in 20 MHz band
Thermal noise level	-174 dBm/Hz
MC and SC noise figure	9 dB in 20 MHz band
Antenna gain	17 dBi
Antenna height	MBS: 25m; User: 1.5m
Average building height	21 m
Channel model	3D model from [3]
Propagation type	MBS to User: LoS; Interferer links: NLoS
Number of interferer links	17
Number of simulations	1000 simulations, each with a random deployment of User for each cell
Training model	Bayesian Regularization (BR)
Number of hidden layers	8
Number of delays (d)	1
Training data set	75% data
Validation data set	15% data
Testing data set	15% data

which provides an average RE (average of the seven days samples for all the cells in the scenario) of 0.132.

$$RE = \frac{\text{Real data} - \text{Predicted data}}{\text{Real data}} \tag{5.6}$$

All the assumptions related to the evaluation are summarised in Table 5.2.

5.2.5 Results

In this section, we present the obtained results in a comparative manner for different approaches described in Section 5.2.4. We discuss the results in three different phases; i) peak-hour analysis (i.e., time slot with the highest load), ii) 24 hour analysis (average of seven days converted into one day), and iii) machine learning-based traffic prediction and BW allocation for seven days.

Note that, in this work, we are interested in testing our proposed dynamic approach in a 5G-like environment, and thus, we scaled up the collected LTE throughput requirements into a 5G scenario, considering that the traffic shows the same behaviour as in LTE. In [60], ITU-T predicted that, compared to recent LTE networks, data rate requirements in 5G will be 10-folded, whereas, according to NTT DOCOMO, the increase ratio will be closer to 100 folds [141]. Therefore, we considered different numbers, i.e., from 10 to 120, as a Scaling Factor (SF), with which we multiply the observed LTE carried traffic to scale it in order to resemble a challenging 5G scenario.

5.2.5.1 Peak hour analysis

Initially, we select the peak hour (i.e., the time slot when a cell experiences the maximum load during a day) to evaluate our proposed access-aware dynamic BW allocation approach for different values of the SF. In this peak hour scenario, we compare the performance of the proposed approach with two static approaches, as described in Section 5.2.4.

Analyzing the available data set, we found that most cells experience the highest load of the day around 23:15:00h, as shown in Figure 5.7. Figure 5.10 presents the number of unsatisfied cells (i.e., the bandwidth assignment could not meet the actual requirements, as read from the network traces, and scaled by the SF) during the peak hour (23:15:00h) for different levels of SF. Clearly, our proposed dynamic approach outperforms the two traditional static approaches by serving more cells (≈ 10), while utilizing the same level of system BW (i.e., 100 MHz). At a very high level of SF (i.e., 90 to 120), the number of unsatisfied cells for Static approach 2 and dynamic approach is very close. This is due to the very high requirements, which are really difficult to meet by nowadays LTE deployments

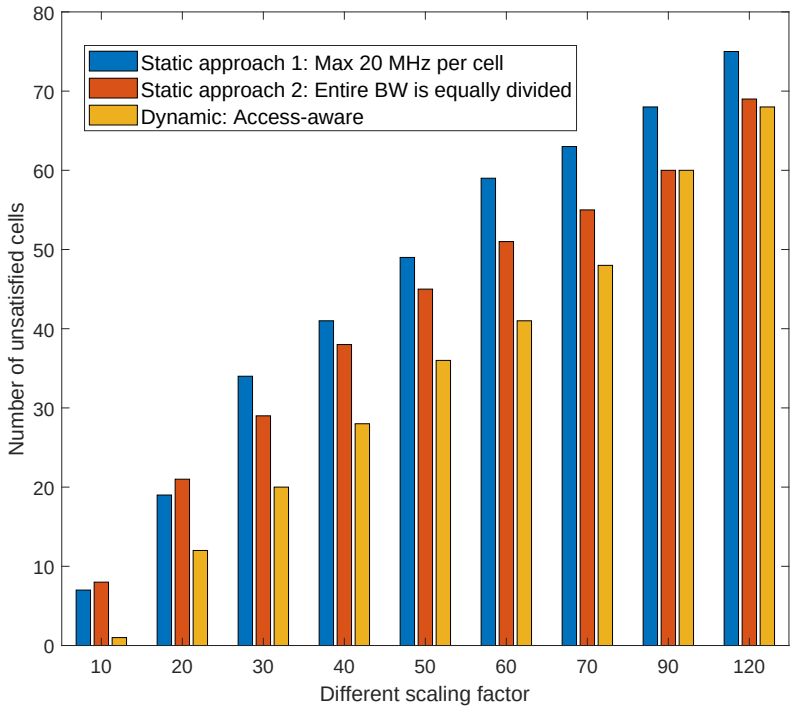

Figure 5.10: Number of unsatisfied cells with different levels of SF.

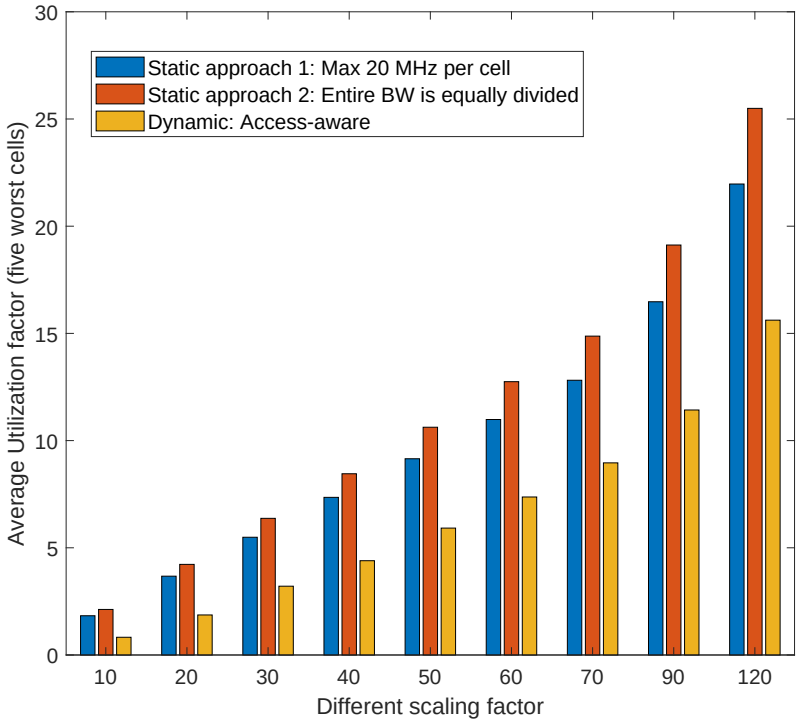

Figure 5.11: Average UF of the worst five cells with maximum load for different levels of SF.

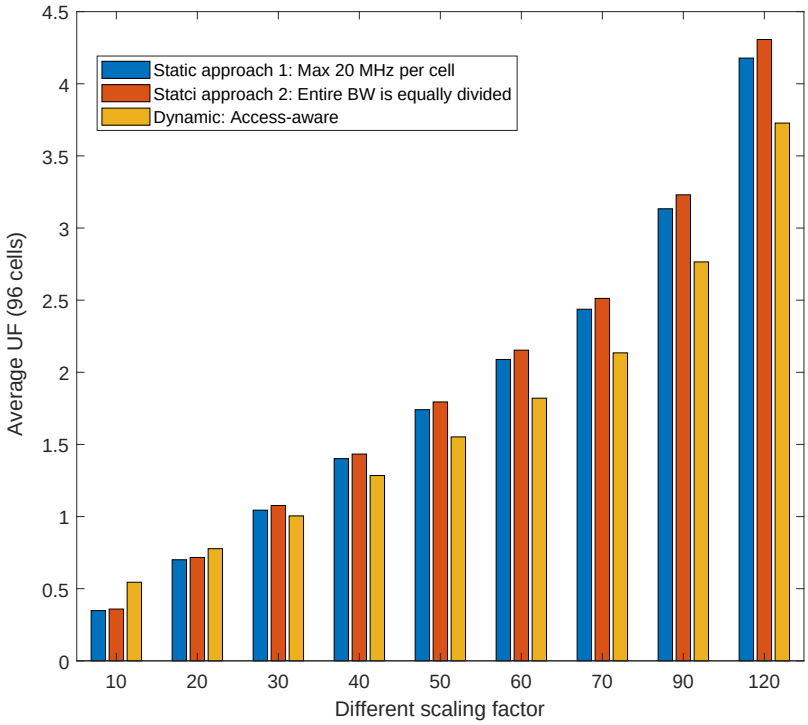

Figure 5.12: Average UF of 96 cells for different levels of SF.

such as the one used as the basis of our model, designed to support a lighter load. In that extreme, the deployment of a denser network (i.e., more BSs) will be a necessity. On the other hand, in terms of the Utilization Factor (UF), which we define as the ratio between required capacity vs. the achieved capacity for a cell, the proposed dynamic access-aware approach performs better for any level of SF. Its performance improves for higher SFs, being it especially efficient for large loads (e.g., SFs 90 and 120). To ensure a fair distribution of BW and to maximize the spectral efficiency, it is better to keep the UF value as close to 1 as possible. Values smaller than 1 mean that the network is overprovisioned and frequency resources are wasted, while values larger than 1 imply that capacity requirements are not fully met. Figure 5.11, depicts the average UF of the five cells with the highest load (identified from the previously analyzed unsatisfied cells) for different levels of SFs, which shows a fairer and better distribution in a dynamic access-aware BW allocation approach, compared to the static approaches. The results presented in Figure 5.11 also show that, utilizing the same level of system BW, the dynamic access-aware approach, is capable of serving more traffic (i.e., UF closer to 1) compared to the static approaches.

Figure 5.12 shows the UF averaged over the 96 cells. The results present a similar trend as in Figure 5.11, i.e., the proposed dynamic access-aware approach performs better (i.e., UF values closer to 1) than the other two static approaches, and for higher values

of SF the performance gain becomes more evident as congestion increases (i.e., UF values higher than 1). Additionally, from Figure 5.12 we can conclude that SF between 20 - 30 consists in the scenarios where we can actually meet the maximum capacity requirements (UF $\leq$ 1) utilizing the available system BW (i.e., 100 MHz per MBS). For larger SF, the studied deployment, designed to support current traffic demands, becomes unable to serve the projected requirements (UF $\geq$ 1). In other words, with a scaling factor lower than 30, the studied network has the potential to carry the scaled traffic and, therefore, remains useful for the study of a future 5G scenario. In contrast, when scaling the traffic x120, a network 4 times denser would be needed (i.e., 4 times the number of deployed MBSs). Therefore, in the subsequent sections, we use three levels of SF, i.e., 20, 25 and 30, to present the evaluation results.

5.2.5.2 24 hour analysis

In this subsection, we analyze the results in a 24 hour scenario, utilizing the average results of seven days. In Figure 5.13 we present the average UF of 96 cells, varying with the time of the day. As mentioned earlier, the UF value closer to value 1 states that available BW is more efficiently used. As depicted in Figure 5.13 for each SF (i.e., 20, 25 and 30), the proposed dynamic access-aware approach uses less resources than the other two static approaches to serve the same amount of traffic during the off-pick hour. On the other hand, during the pick-hour and for the higher value of SF (i.e., 30), it is shown that, the dynamic access-aware solution serves more traffic (i.e., UF closer to 1) in a congested scenario. Therefore, utilizing the same level of system BW, and employing our proposed dynamic distribution, the congestion (i.e., UF > 1) period is 7% shorter (in a 24 hour scenario for SF = 30), compared to the Static approach 2.

Figure 5.14 depicts the number of unsatisfied cells, (i.e., those cells that could not reach the required capacity), during the day for the different approaches. The dynamic access-aware approach tends to satisfy a higher number of cells given the BW limitations for a MBS. For instance, for SF 20, during off-peak hour, i.e., low load hour (03:00:00h to 08:15:00h), the dynamic access-aware approach is able to serve all the 96 cells without any congestion. However, both the static approaches have a few number of congested cells. On the other hand, during the peak hours (i.e., 20:00:00h to 23:45:00h), dynamic access-aware approach satisfies 9% - 12% more cells compared to the static approaches. With the highest SF of 30, the performance gain of the dynamic access-aware approach is even higher (i.e., 10% - 15% more satisfied cells) compared to the two static approaches. Therefore, the proposed dynamic access-aware BW allocation approach appears as a better solution for the congestion

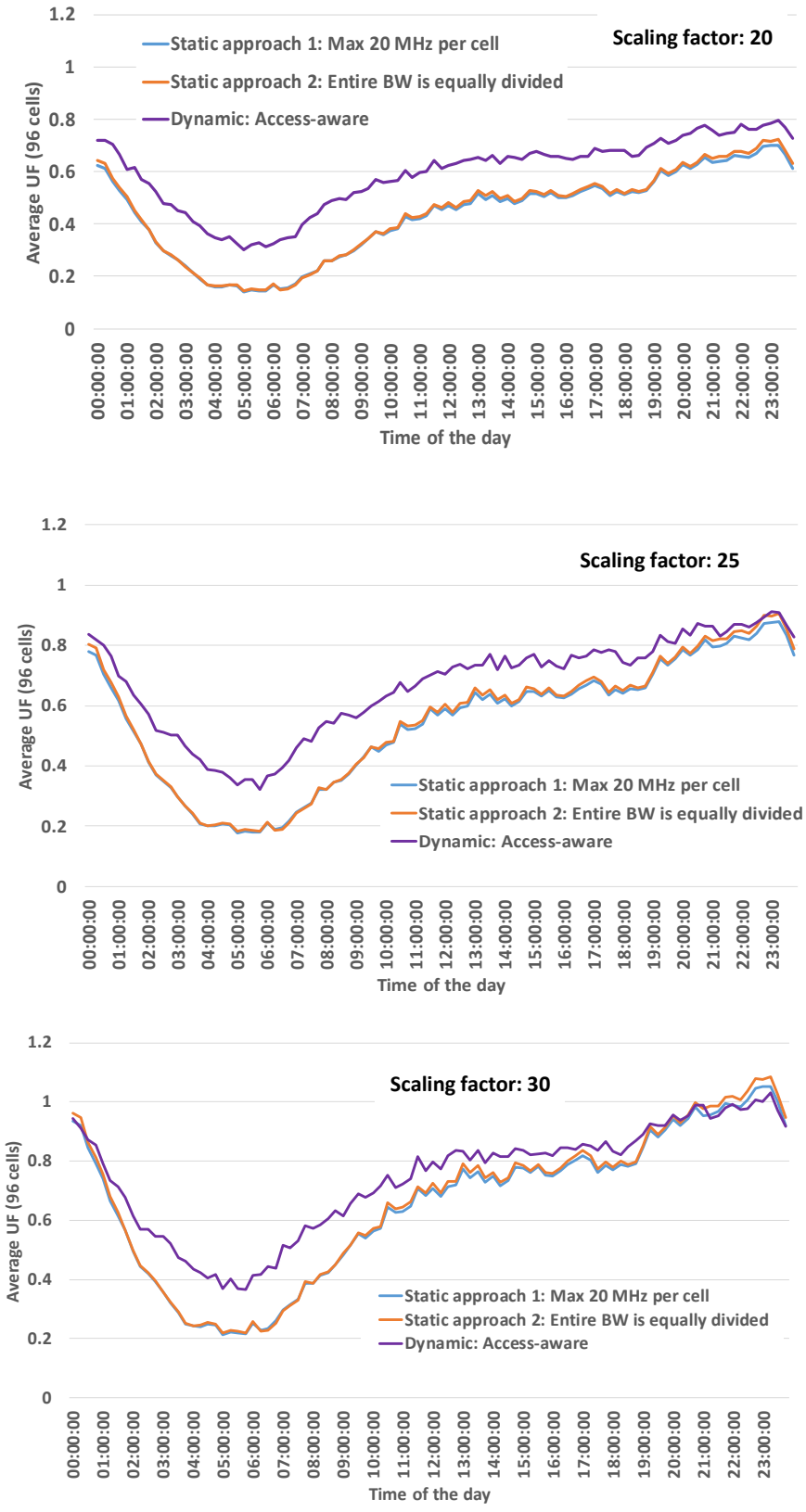

Figure 5.13: Average UF of 96 cells for different approaches.

problem in the network.

At this point, we also want to observe how much system BW the proposed dynamic access-aware BW allocation approach can save by assigning the necessary amount of BW following the access networks' requirements. In Figure 5.15, the total BW assigned (average of 24 hours) to each of the studied 21 MBS is depicted for the three different approaches. For Static approach 1, the assigned BW for a MBS depends on the number of cells it has (with a maximum of 20 MHz per cell), remaining agnostic to the access networks' current requirements, and thus, being independent of the SF. On the other hand, for Static approach 2, the total assigned BW for each MBS is 100 MHz, independent of the number of cells, or the dynamic requirements of the access network. On the contrary, the total BW assigned for a MBS in the dynamic access-aware approach is a total reflection of the throughput requirements of the access networks and, to a lesser extent, of the number of cells of a MBS, given that all cells are assigned at least 5 MHz. Moreover, the total assigned BW changes for each cell and MBS, if required, depending on the SF. Thus, in Figure 5.15, the total BW for a MBS varies for different levels of SF, contrary to the static approaches, where it remains unchanged and agnostic to the current network condition. Additionally, as depicted in Figure 5.15, for the highest SF studied (i.e., 30), the dynamic access-aware approach can save from 5 MHz (i.e., MBS 17) to 60 MHz (i.e., MBS 10) of system BW, compared to Static approach 1, whereas, the number can be between 10 MHz (i.e., MBS 1) to 90 MHz (i.e., MBS 21) compared to Static approach 2.

5.2.5.3 Machine learning based traffic prediction and BW allocation

In this subsection, we focus on the ML predicted throughput requirements and the variation in the BW allocations based on the results of the predictions.

Thus, this subsection compares the quality of the BW assignments when the system has perfect prior knowledge of the traffic demands, with the case where the algorithm relies on traffic predictions produced by the neural network described in Section 5.2.4. In Figure 5.16, varying with the time of seven days, we present the average (i.e., an average of 96 cells) real traffic requirements as extracted from the traces, and the average requirements predicted by the trained NARNET for the same time instant. As depicted in Figure 5.16, our selected ML approach provides a prediction showing mean RE = 0.132, and follows the behavior of the real trace. However, it is also observed that predicted traces are slightly higher than the real traces, considering the average of 96 cells. This difference has a clear reflection on the BW assignment, because, in our proposed dynamic access-aware BW allocation approach, BW is allocated according to the throughput requirements.

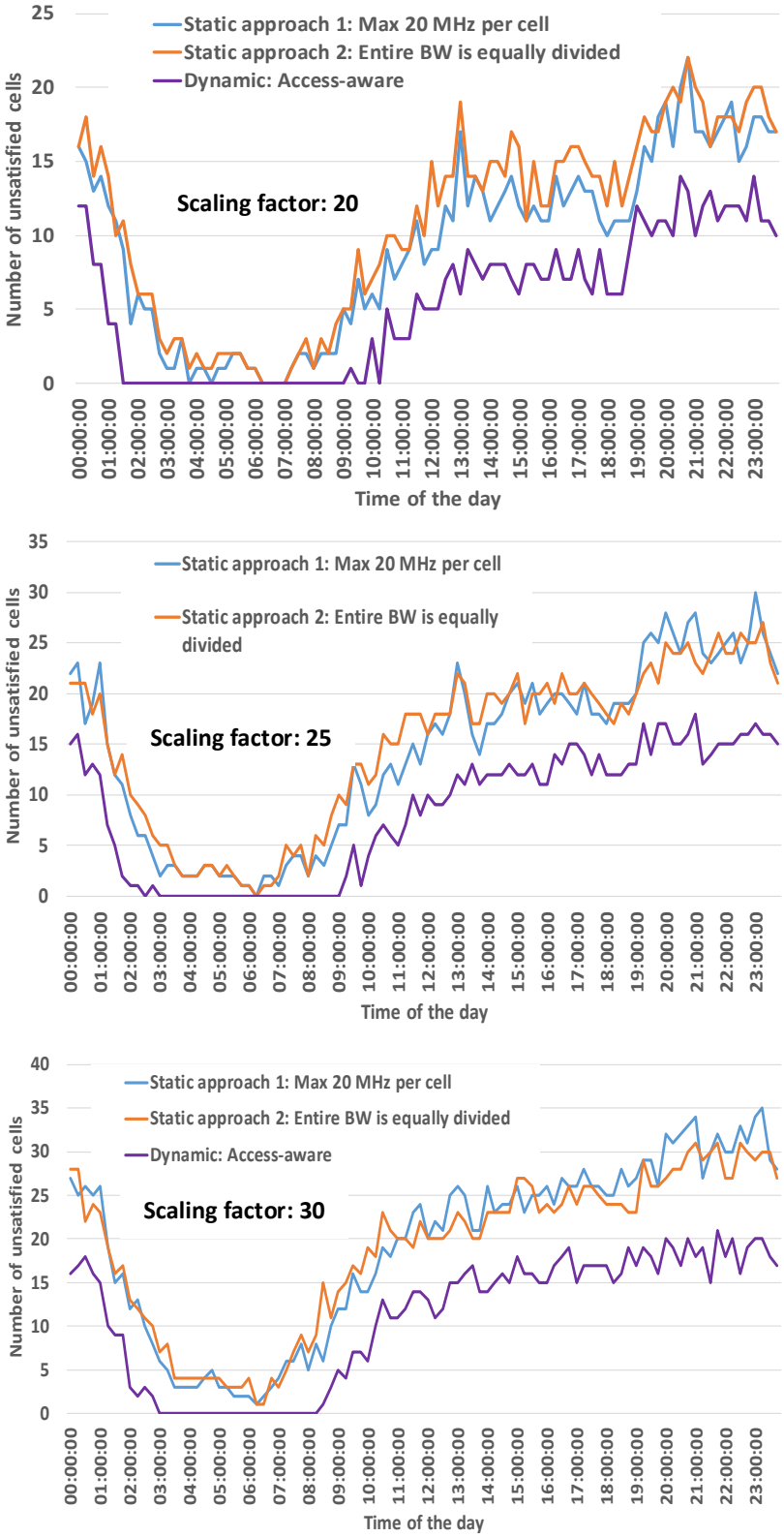

Figure 5.14: Number of unsatisfied cells during the day for different approaches.

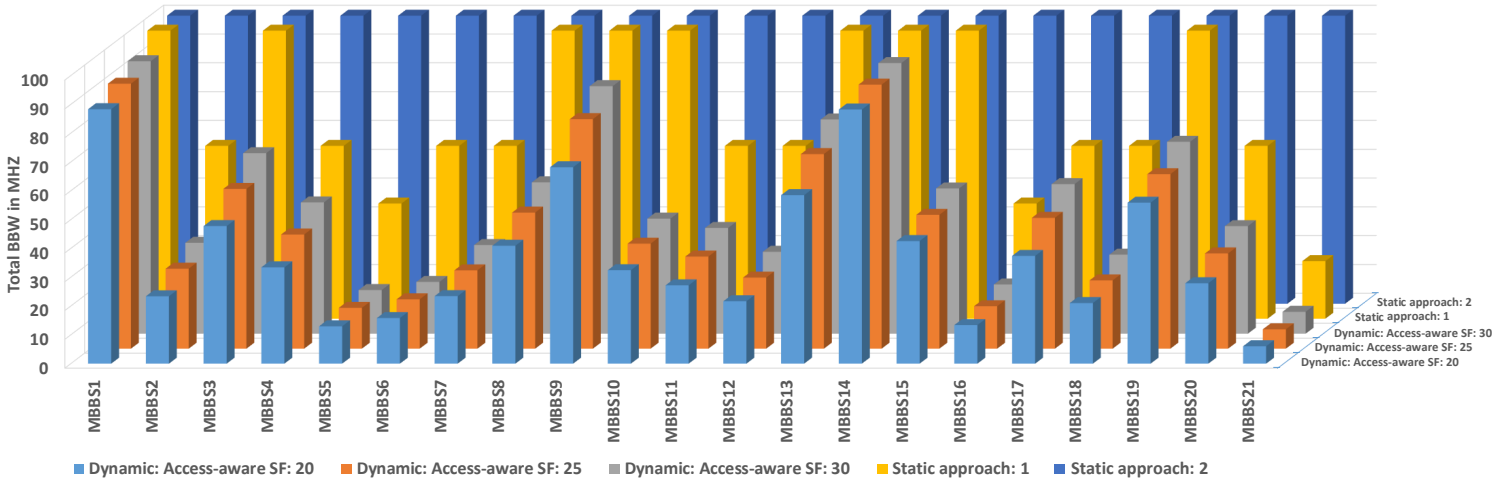

Figure 5.15: Total allocated BW to each MBS for different BW allocation approaches.

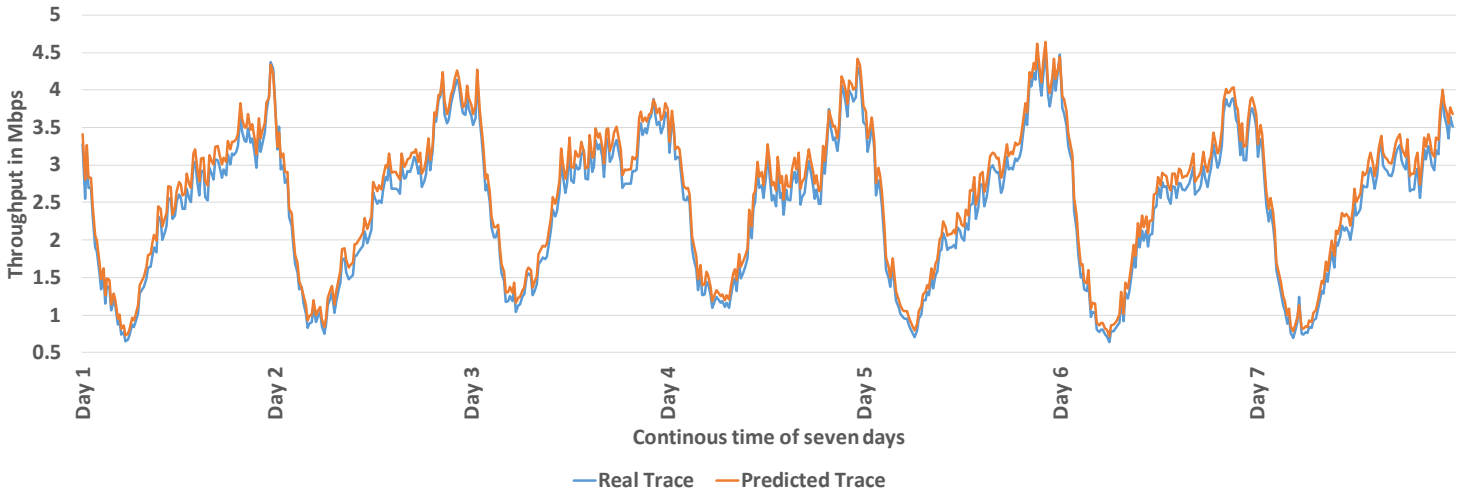

Figure 5.16: Real and predicted average (over 96 cells) throughput requirements for seven days.

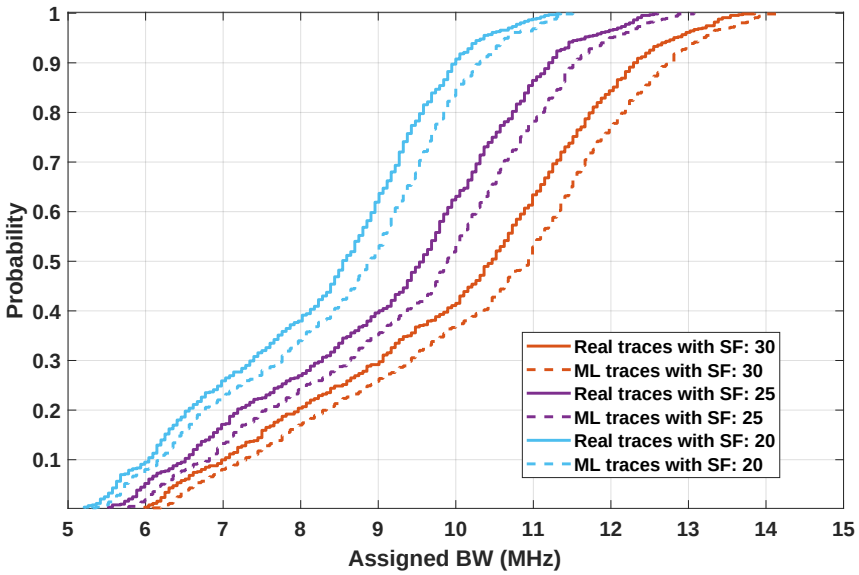

Figure 5.17: CDF of average (over 96 cells) assigned BW for perfect knowledge of real traffic (Real) and ML predicted traffic, using different scaling factors (SF).

Figure 5.17 depicts the Cumulative Distribution Function (CDF) plot of the average (over 96 cells) assigned BW (in MHz) to a cell during seven days following real and predicted traces for different levels of SF. As discussed earlier, the predicted traces have usually shown higher values than the real traces, hence, the assigned BW based on those predictions is also slightly higher than strictly needed in practice. For instance, depicted in Figure 5.17, with SF of 25, a cell is satisfied with 10 MHz or less, on average, 62% of the time, whereas, following ML predictions, this is reduced to 52% of the time. A similar trend follows the results observed for SF values of 20 and 30, where cells are assigned a slightly higher BW when the mechanism is based on predictions instead of having a perfect knowledge of future traffic. In practice, this conservative behavior would be more robust in front of unexpected increments in traffic, at the cost of less adjusted UF.

In Figure 5.18, we present an analysis of UF for the four BW assignment approaches: i) static approach 1, ii) static approach 2, iii) ideal dynamic approach assuming perfect knowledge of future demands (i.e., following real traces), and iv) dynamic approach following ML predictions. Note that for all the cases, UF is computed with respect to the real requirements, i.e., for iv), UF is the ratio between real throughput requirements vs. the achievable capacity of the network configured based on traffic predictions. Figure 5.18 presents the CDF plot of average UF (over 96 cells) during seven days (i.e., 672 samples: 7 days with 15 minutes granularity) for different SFs (i.e., 20, 25 and 30). As depicted in Figure 5.15, due

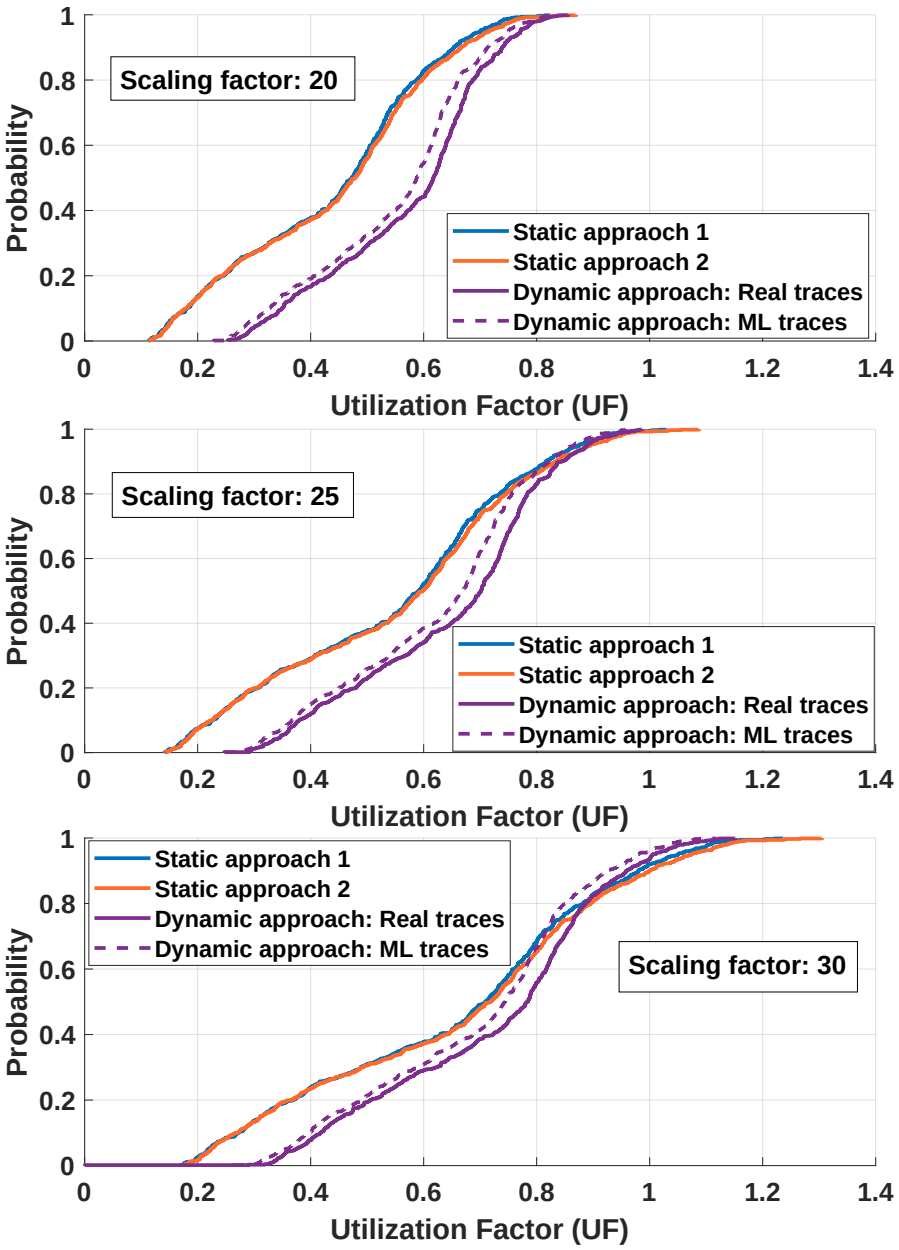

Figure 5.18: CDF of average (over 96 cells) UF for perfect knowledge of real traffic (Real); and ML predicted traffic, using different scaling factors (SF).

to slightly higher throughput predictions, the predicted BW allocation is also slightly higher than the ideal access-aware dynamic BW allocation approach, and thus, UF (Real traces) in Figure 5.18 for different levels of SF are slightly better (i.e., closer to value 1) compared to the ML approach. On the other hand, it is also observed that ML-based BW assignment results into better UF (i.e., closer to value 1) compared to the two static approaches during satisfied (UF < 1) and unsatisfied (UF > 1) period.

As shown in Figure 5.18, for the dynamic approach following the real traces (iii), and the dynamic approach following the predicted traces (iv), UF values remain acceptably close, showing mean RE = 0.0485 for SF = 20, mean RE = 0.0463 for SF = 25, and mean RE = 0.0475 for high SF = 30. Therefore, in terms of utilized resources, the efficiency of a solution based on traffic predictions is similar to the performance obtained if perfect knowledge of future traffic demands were possible. Hence, the predictions from a well-trained NARNET, which is performed utilizing a batch of real data set, can be very useful to predict the future requirements and perform an efficient allocation of BW to serve the real traffic in the network. Moreover, as depicted in Figure 5.10, 5.11, 5.13, and 5.18 for different levels of SF, our proposed dynamic approach always performs better than the static approaches.

As a limitation of this work, we believe that the average UF_Real and UF_ML are not very remarkable (i.e., should be more close to value 1), this is because, during the off-peak hour (03:00:00h to 08:15:00h) the required throughput is usually low compared to the configured capacity even utilizing the minimum offered BW (i.e., 5 MHz) considered. This can be an area of improvement for this work to look for more fine-grained channel BW for such future potential mid-band RAT (i.e., Sub-6 GHz). Also note that, limited by the granularity of collected traces, the proposed access-aware dynamic spectrum allocation technique triggers every 15 minutes. We believe, a shorter Reporting Output Period (ROP) to collect the traces can make the proposed approach more dynamic and spectral-efficient, without increasing complexity, since the application of the NARNET is computationally cheap, once trained.

5.3 Summary and future work

Expected capacity requirements of future wireless networks continue to challenge the traditional approaches of frequency resource allocation. Moreover, the access network requirements vary dynamically, which calls for a dynamic allocation of the frequency resources to ensure the efficient usage of precious resources.

The possibility that both BH and access networks share spectrum resources, brings the need of cooperative spectrum sharing. Therefore, Section 5.1 identifies three types of poten-

tial links in future RAN competing for the same spectrum from a joint access-backhaul mechanism perspective. Presented comparative results show that in terms of spectral efficiency, fairness and achievable throughput, access-aware cooperative spectrum sharing techniques outperform the traditional approaches, where the spectrum allocation does not consider access-BH awareness. Therefore, resources assigned to both BH and access should follow the varying demands of the access. In Section 5.2 we have presented a ML technique to predict those variations and apply this knowledge to assign resources to the access network. As a consequence, assignments to BH according to those predicted access requirements should follow access-aware approaches, i.e., SSA-2(a), SSA-3(a), SSA-5(a), as suggested in Section 5.1, which we leave for future studies. However, we believe that the characterization and understanding of different spectrum sharing alternatives provided in Section 5.1 are useful for the development of optimization algorithms that will make the most of those scarce resources.

In Section 5.2, the presented access-aware dynamic BW allocation approach follows the current requirements of each cell and allocates the required BW to serve the carried traffic. The evaluation results show that, with the dynamic access-aware approach, the frequency resources are used more efficiently, i.e., UF closer to 1. Additionally, such intelligent allocation can serve more traffic, specially during peak hours. Utilizing the same level of system BW, the dynamic allocation approach recovers 9% - 12% cells from congestion compared to the traditional static approaches. We have presented the evaluation results by scaling requirements of a real LTE network into a 5G-like scenario, and concluded that, with the higher values of scaling factor, the performance gain of the dynamic access-aware approach is more tangible in terms of the number of congested cells.

However, having the perfect network knowledge (i.e., real traces) is a challenge, which cannot be taken for granted. Therefore, in this work, we have presented a NARNET-based ML technique to predict the access network requirements with an acceptable level of error (i.e., RE: 0.132). Presented comparative results show that, in terms of resource utilization and BW allocation, following a NARNET-based prediction, the dynamic access-aware BW allocation approach performs very close to the results obtained with perfect network knowledge.

Chapter 6

Conclusions and future work

Joint access-backhaul mechanism is one of the key enablers for 5G deployments. To meet the aggressive demand of 5G, efficient usage of scarce resources is unavoidable, and joint access-backhaul operation can ensure the programmability that 5G requires. Additionally, the design of 5G backhaul networks will be very challenging, since it will be complex, heterogeneous, very high data demanding, and probably the most cost-dominant element of the 5G architecture. The intelligent implementation of this kind of joint mechanism can meet these challenges to ensure the transition towards 5G.

6.1 Conclusions

In this thesis, utilizing the programmability of joint access and backhaul networks, we have identified three very attractive areas, which address the cost and capacity limitation of backhaul networks. Additionally, the presented novel ideas contribute to the development of the cell-less concept with the study of different centralized mechanisms for resource sharing in the access network (among different cells), backhaul/fronthaul network (among different links), and among each other (access and backhaul).

Chapter 2, presents a detailed literature study on joint access-backhaul mechanisms. This chapter showcases potential approaches to perform a joint mechanism, and additionally, a few related FP7, H2020 and 5GPPP projects and their vision/outcomes are summarized. Literature study is a continuous process, which we believe, will bring more opportunities and challenges on the table, opening further research streams to approach the joint access-backhaul concept. Our contribution [C1] summarizes the state-of-the-art in joint access-backhaul mechanism, and additionally, proposes a research direction arguing that mobility management in 5G access network should be backhaul-aware, thus validating the joint mech-

anism.

In Chapter 3, we propose a novel technique, the access-aware backhaul optimization, where backhaul resources are assigned according to the current requirement on the access network ensuring the best utilization of precious resources. The presented results are used as an initial evaluation, showing that the proposed concept of awareness between access and backhaul network ensures efficient utilization of the scarce resources. Discussions in this chapter motivate and drive the subsequent chapters, where more developed research ideas to implement the concept of joint access-backhaul mechanism are proposed and evaluated. Our contribution [C2] reflects the work presented in Chapter 3.

In Chapter 4, we analyze cost optimization of FH links based on functional splits at the PHY layer, arguing that the design of APs and their functionalities have a large impact on the FH network's OPEX and TCO. Moreover, the stated challenge of capacity-hungry FH network and their economical impact can be tackled utilizing different functional splits at the PHY layer. The presented results show that the heterogeneous combination of functional splits at the PHY layer provides a potential solution to serve RAN and FH network in a cost and capacity-limited scenario. It is also shown that different approaches can be adopted while minimizing the deployment cost or maximizing the centralization, according to operators' priority. Our contribution [C4] discusses the impact of functional splits on OPEX, whereas, [J1] provides a solution to minimize the TCO utilizing different combinations of functional splits at the PHY layer. Both the contributions include the opportunity to prioritize the centralization gain in a capacity-limited scenario.

In Chapter 5, joint spectrum sharing, takes the challenge of limited spectrum, considering the high requirements of a 5G network. Traditional approaches are not efficient enough to take this challenge and, thus, we propose and test a joint spectrum sharing and dynamic spectrum allocation technique that ensures the best utilization of spectrum, and maximizes the resource efficiency of the network. Utilizing the numerical results presented in Section 5.1, our contribution [C3] concludes that the access-aware joint (access and backhaul) spectrum allocation technique ensures the efficient usage of scarce resources. In Section 5.2, we address the limitation of the work presented in Section 5.1, which is a static solution. Therefore, in Section 5.2, we provide an access-aware dynamic spectrum allocation technique only in the access network. Note that we leave the dynamic access-aware joint (access and backhaul) spectrum allocation technique for our future work. Additionally in Section 5.2, machine learning-based traffic prediction and BW allocation following the dynamic requirements are presented. The discussed results show that adaptation techniques can be resource-efficient in a programmable network for BW allocation to different cells following the dynamic behaviour of the traffic, provided that such behavior is predicted, for example, by a well-trained ML

algorithm. Our contribution [J2] mirrors the discussions presented in Section 5.2.

6.2 Future research directions

The presented works bring opportunities for future research. Cost optimization of FH network allows further research utilizing the functional splits within higher layers, i.e., high-MAC, low-MAC (cf. Figure 4.2), and among different layers, i.e., PCDP-RLC, as presented in [69] and discussed in Section 4.1. With this, FH capacity requirements will be further decreased, opening further the set of available technologies capable of serving the needs of a fronthaul, which, in turn, will have a large impact on OPEX and TCO. In that case, subsequent compromise of centralization gain will need further studies. Additionally, we note that the presented brute force algorithm might not be efficient for a large scale deployment scenario, where we have the opportunity to develop more advanced ML-based algorithms to automate the PHY layer splitting decision. Moreover, the presented idea proposes a static deployment approach to find the best combination of different splits according to different optimization objectives (i.e., cost minimization and centralization gain maximization). We believe, there is a great opportunity to study the presented idea for a dynamic approach, where an existing network can change the PHY layer splits dynamically according to the traffic demand, OPEX minimization or required level of centralization. To study this, the feasibility of adapting dynamic changes in the PHY layer splits also requires further investigation.

Our presented dynamic spectrum allocation technique studies the allocation of bandwidth to the cells, however, the frequency assignment problem was not studied. This brings an opportunity for further research, where ML-based dynamic frequency assignments following the dynamic traffic requirements can bring new opportunities to the frequency planning for future networks [142]. Additionally, the presented dynamic spectrum allocation technique (Section 5.2) has not been tested within a complete joint access-backhaul mechanism. On the other hand, cooperative spectrum sharing (Section 5.1) in access and backhaul has been presented in a static spectrum sharing approach. Further research on dynamic spectrum allocation in a joint access-backhaul perspective should be very interesting since it will merge both the presented spectrum sharing approaches.

Additionally, as motioned in state-of-the-art section, a few other approaches, i.e., joint coding, joint load balancing, joint interference management, joint energy optimization, etc. have the potential to facilitate joint access-backhaul mechanism.

Throughout the years, the hardware-based network architecture has been evolving to meet traffic demands and to create new opportunities in mobile communications. On the

other hand, this hardware-based network architecture also limits the network performance imposing new challenges. Collaboration of software and hardware becomes necessary to explore the new opportunities of wireless communications. With this vision, SDN is being considered as one of the prime enablers for 5G networks. Employing the SDN paradigm in wireless networks can boost the network performance, while making the network more flexible, reliable and scalable. With the separation of control and data planes, SDN enhances the opportunity to centralize the control section and to optimize the network centrally. This approach makes joint access-backhaul mechanism more realistic and efficient. With an SDN controller, network can be more adaptive with the complex commutation and fast decision making, which is very crucial for joint access-backhaul mechanism and the presented use cases in this thesis.

3GPP release 16 has already standardized integrated access backhaul (IAB) to provide multi-hop backhauling, which challenges the traditional relay approach only capable of two hops relaying. Standardized IAB opens up more opportunities to implement joint access-backhaul mechanisms.

Bibliography

[1] M. Jaber, D. Owens, M. A. Imran, R. Tafazolli, and A. Tukmanov, "A joint backhaul and ran perspective on the benefits of centralised ran functions," in *IEEE ICC*, May 2016, pp. 226–231.

[2] METIS-II, "Performance evaluation framework," *Deliverable D2.1*, no. January, p. 10, 2016. [Online]. Available: https://metis-ii.5g-ppp.eu/wp-content/uploads/deliverables/METIS-II_D2.1_v1.0.pdf

[3] 5G-PPP, "5G PPP use cases and performance evaluation models," *Living document on*, no. March, p. 10, 2016. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-use-cases-and-performance-evaluation-modeling_v1.0.pdf

[4] T. Mshvidobadze, "Evolution mobile wireless communication and LTE networks," *Appl. Inf. Commun. Technol. (AICT), 2012 6th Int. Conf.*, pp. 1–7, 2012.

[5] Ericsson, "Ericsson Mobility Report (June 2020)," no. June, p. 36, June, 2020. [Online]. Available: https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf

[6] I. F. Akyildiz, S. Nie, S.-C. Lin, and M. Chandrasekaran, "5g roadmap: 10 key enabling technologies," *Computer Networks*, vol. 106, pp. 17–48, 2016.

[7] F. Testa, F. Cavaliere, and R. Sabella, "Future Generation of Wireless Communication Systems: requirements and open issues," *Microwave Photonics workshop at ECOC 2015, Valencia, Spain.*

[8] G. K. Chang and L. Cheng, "Fiber-wireless integration for future mobile communications," in *2017 IEEE Radio and Wireless Symposium (RWS)*, Jan 2017, pp. 16–18.

[9] E. Pateromichelakis, A. Maeder, A. D. Domenico, R. Fritzsche, P. D. Kerret, and J. Bartelt, "Joint RAN / Backhaul Optimization in Centralized 5G," pp. 386–390, 2015.

[10] Ericsson, "5G Radio Access What is 5G ?" *White Pap.*, no. February, p. 10, 2015. [Online]. Available: http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf

[11] J.-i. Kani, S. Kuwano, and J. Terada, "Options for future mobile backhaul and fronthaul," *Optical Fiber Technology*, vol. 26, pp. 42–49, 2015.

[12] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "5G Backhaul Challenges and Emerging Research Directions: A Survey," *IEEE Access*, vol. 4, pp. 1743–1766, 2016.

[13] J. Bartelt, G. Fettweis, D. Wübben, M. Boldi, and B. Melis, "Heterogeneous backhaul for cloud-Based mobile networks," *IEEE VTC*, 2013.

[14] A. H. Jafari, D. López-Pérez, H. Song, H. Claussen, L. Ho, and J. Zhang, "Small cell backhaul: challenges and prospective solutions," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, p. 206, 2015.

[15] A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things you should know about fronthaul," *Journal of Lightwave Technology*, vol. 33, no. 5, pp. 1077–1083, 2015.

[16] F. Cavaliere, P. Iovanna, J. Mangues-Bafalluy, J. Baranda, J. Núñez-Martínez, K.-Y. Lin, H.-W. Chang, P. Chanclou, P. Farkas, J. Gomes *et al.*, "Towards a unified fronthaul-backhaul data plane for 5g the 5g-crosshaul project approach," *Computer Standards & Interfaces*, vol. 51, pp. 56–62, 2017.

[17] IEEE Standard Association, "IEEE 802.3bm-2015 - IEEE Standard for Ethernet - Amendment 3: Physical Layer Specifications and Management Parameters for 40 Gb/s and 100 Gb/s Operation over Fiber Optic Cables," *IEEE 802.3bm-2015*, 27 March, 2015.

[18] 5G-XHaul, "System Architecture Definition," *Deliverable D2.2*, no. January, p. 73, 2016. [Online]. Available: https://www.5g-xhaul-project.eu/download/5G-XHaul_D2_2.pdf

[19] European Telecommunications Standards Institute (ETSI), "5g wireless backhaul/X-Haul," *GR mWT 012*, November, 2018. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/mWT/001_099/012/01.01.01_60/gr_mWT012v010101p.pdf

[20] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "A multiple attribute user-centric backhaul provisioning scheme using distributed son," in *IEEE GLOBECOM*. IEEE, 2016, pp. 1–6.

[21] NGMN-Alliance, "5g white paper," *Next generation mobile networks, white paper*, 2015.

[22] O. Marinchenco, N. Chayat, M. Goldhamer, A. Burr, A. Papadogiannis, M. Dohler, M. Payaro, C. Oestges, N. Khan, and M. Ware, "BuNGee project overview," *IEEE COMCAS*, 2011.

[23] C. Alvarion, P. T. Cyfrowa, and T. Siklu, "Broadband radio access networks (bran); very high capacity density bwa networks; protocols," *ETSI, Valbonne, France, Tech. Rep. TC BRAN, TR*, vol. 101, no. 589, p. V1, 2012.

[24] C. Dehos, J. L. González, A. D. Domenico, D. Kténas, and L. Dussopt, "Millimeter-wave access and backhauling: the solution to the exponential data traffic increase in 5g mobile communications systems?" *IEEE Communications Magazine*, vol. 52, no. 9, pp. 88–95, September 2014.

[25] S. González, A. Oliva, X. Costa-Pérez, A. Di Giglio, F. Cavaliere, T. Deiß, X. Li, and A. Mourad, "5g-crosshaul: An sdn/nfv control and data plane architecture for the 5g integrated fronthaul/backhaul," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 9, pp. 1196–1205, 2016.

[26] A. Ravanshid, P. Rost, D. S. Michalopoulos, V. V. Phan, H. Bakker, D. Aziz, S. Tayade, H. D. Schotten, S. Wong, and O. Holland, "Multi-connectivity functional architectures in 5g," in *2016 IEEE International Conference on Communications Workshops (ICC)*, May 2016, pp. 187–192.

[27] K. Samdanis, R. Shrivastava, A. Prasad, D. Grace, and X. Costa-Perez, "Td-lte virtual cells: An sdn architecture for user-centric multi-enb elastic resource management," *Computer Communications*, vol. 83, pp. 1–15, 2016.

[28] 5G-XHaul project Homepage [Online], http://www.5g-xhaul-project.eu/index.html, accessed: 2017-10-25.

[29] J. Gamboa and I. Demirkol, "Softwarized lte self-backhauling solution and its evaluation," in *2018 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2018, pp. 1–6.

[30] F. Tonini, C. Raffaelli, B. M. Khorsandi, S. Bjornstad, and R. Veisllari, "Converged fronthaul/backhaul based on integrated hybrid optical networks," in *Asia Communications and Photonics Conference*. Optical Society of America, 2018, pp. Su1E–1.

[31] M. Grandi, D. Camps-Mur, A. Betzler, J. J. Aleixendri, and M. Catalan-Cid, "Swam: Sdn-based wi-fi small cells with joint access-backhaul and multi-tenant capabilities," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. IEEE, 2018, pp. 1–2.

[32] A. Betzler, D. Camps-Mur, E. Garcia-Villegas, I. Demirkol, and J. J. Aleixendri, "Sodalite: Sdn wireless backhauling for dense 4g/5g small cell networks," *IEEE Transactions on Network and Service Management*, vol. 16, no. 4, pp. 1709–1723, 2019.

[33] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "On the joint optimisation of radio access and backhaul networks," in *2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*. IEEE, 2017, pp. 1–5.

[34] X. Wang, C. Cavdar, L. Wang, M. Tornatore, Y. Zhao, H. S. Chung, H. H. Lee, S. Park, and B. Mukherjee, "Joint allocation of radio and optical resources in virtualized cloud ran with comp," in *IEEE GLOBECOM*, Dec 2016, pp. 1–6.

[35] A. Maeder, M. Lalam, A. De Domenico, E. Pateromichelakis, D. Wubben, J. Bartelt, R. Fritzsche, and P. Rost, "Towards a flexible functional split for cloud-RAN networks," *EuCNC 2014*, pp. 5–9.

[36] C. J. Bernardos, A. D. Domenico, J. Ortin, P. Rost, and D. Wübben, "Challenges of designing jointly the backhaul and radio access network in a cloud-based mobile network," in *2013 Future Network Mobile Summit*, July 2013, pp. 1–10.

[37] D. Wubben, P. Rost, J. S. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5g signal processing: Flexible centralization through cloud-ran," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 35–44, Nov 2014.

[38] J. Bartelt, P. Rost, D. Wubben, J. Lessmann, B. Melis, and G. Fettweis, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 105–111, October 2015.

[39] J. Bartelt, N. Vucic, D. Camps-Mur, E. Garcia-Villegas, I. Demirkol, A. Fehske, M. Grieger, A. Tzanakaki, J. Gutiérrez, E. Grass, G. Lyberopoulos, and G. Fettweis, "5g transport network requirements for the next generation fronthaul interface," *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 89, May 2017. [Online]. Available: https://doi.org/10.1186/s13638-017-0874-7

[40] D. Wubben, P. Rost, J. S. Partelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5g signal processing: Flexible centralization through cloud-RAN," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 35–44, 2014.

[41] U. Dötsch, M. Doll, H. Mayer, F. Schaich, J. Segel, and P. Sehier, "Quantitative analysis of split base station processing and determination of advantageous architectures for lte," *Bell Labs Technical Journal*, vol. 18, no. 1, pp. 105–128, 2013.

[42] R. A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lähetkangas, and E. Tirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wirel. Commun.*, vol. 22, no. 5, pp. 83–89, 2015.

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

[44] D. Wang, E. Katranaras, A. Quddus, N. Sapountzis, L. Cominardi, F. C. Kuo, P. Rost, C. J. Bernardos, and I. Berberana, "SDN-based joint backhaul and access design for efficient network layer operations," in *EuCNC 2015*, pp. 214–218.

[45] O. Dhifallah, H. Dahrouj, T. Y. Al-Naffouri, and M. S. Alouini, "Joint hybrid backhaul and access links design in cloud-radio access networks," *IEEE VTC 2015*.

[46] M. Shariat, E. Pateromichelakis, A. U. Quddus, and R. Tafazolli, "Joint TDD Backhaul and Access Optimization in Dense Small-Cell Networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5288–5299, 2015.

[47] M. Jaber, M. A. Imran, R. Tafazolli, and A. Tukmanov, "A distributed son-based user-centric backhaul provisioning scheme," *IEEE Access*, vol. 4, pp. 2314–2330, 2016.

[48] C. Ran, S. Wang, and C. Wang, "Balancing backhaul load in heterogeneous cloud radio access networks," *IEEE Wirel. Commun.*, vol. 22, no. 3, pp. 42–48, 2015.

[49] X. Xu, W. Saad, X. Zhang, X. Xu, and S. Zhou, "Joint deployment of small cells and wireless backhaul links in next-generation networks," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2250–2253, 2015.

[50] K. M. S. Huq, S. Mumtaz, J. Bachmatiuk, J. Rodriguez, X. Wang, and R. L. Aguiar, "Green hetnet comp: Energy efficiency analysis and optimization," *IEEE TVT*, vol. 64, no. 10, pp. 4670–4683, Oct 2015.

[51] A. Mesodiakaki, F. Adelantado, L. Alonso, and C. Verikoukis, "Energy-efficient context-aware user association for outdoor small cell heterogeneous networks," in *ICC*, June 2014, pp. 1614–1619.

[52] G. Zhang, T. Q. S. Quek, A. Huang, M. Kountouris, and H. Shan, "Backhaul-aware base station association in two-tier heterogeneous cellular networks," in *SPAWC*, June 2015, pp. 390–394.

[53] H. Beyranvand, W. Lim, M. Maier, C. Verikoukis, and J. A. Salehi, "Backhaul-aware user association in fiwi enhanced lte-a heterogeneous networks," *IEEE TWC*, vol. 14, no. 6, pp. 2992–3003, June 2015.

[54] J. Bartelt and G. Fettweis, "Radio-over-Radio: I/Q-stream backhauling for cloud-based networks via millimeter wave links," *GC Wkshps 2013*, pp. 772–777.

[55] A. Sharma, R. K. Ganti, and J. K. Milleth, "Joint backhaul-access analysis of full duplex self-backhauling heterogeneous networks," *IEEE TWC*, vol. 16, no. 3, pp. 1727–1740, March 2017.

[56] S. Singh and J. G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE TWC*, vol. 13, no. 2, pp. 888–901, February 2014.

[57] Y. Lin and W. Yu, "Joint spectrum partition and user association in multi-tier heterogeneous networks," in *48th CISS*, March 2014, pp. 1–6.

[58] U. Siddique and H. Tabassum and E. Hossain, "Spectrum allocation for wireless backhauling of 5g small cells," in *Communications Workshops (ICC), 2016 IEEE International Conference on.* IEEE, 2016, pp. 122–127.

[59] U. Siddique, H. Tabassum, and E. Hossain, "Downlink spectrum allocation for in-band and out-band wireless backhauling of full-duplex small cells," *IEEE Transactions on Communications*, vol. 65, no. 8, pp. 3538–3554, Aug 2017.

[60] ITU, "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond, M Series, Recommendation ITU-R M.2083-0," vol. 0, p. 21, 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf

[61] V. Baños-gonzalez, M. S. Afaqui, E. Lopez-aguilera, and E. Garcia-villegas, "IEEE 802.11 ah: a technology to face the IoT challenge," Sensors 2016, 16(11).

[62] J. Pérez-Romero, J. Sánchez-González, O. Sallent, and R. Agustí, "On learning and exploiting time domain traffic patterns in cellular radio access networks," in *International Conference on Machine Learning and Data Mining in Pattern Recognition*. Springer, 2016, pp. 501–515.

[63] N. Wang, E. Hossain, and V. K. Bhargava, "Backhauling 5g small cells: A radio resource management perspective," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 41–49, 2015.

[64] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 152–160, 2015.

[65] ITU, "Transport network support of IMT-2020/5G - Technical Report - GSTR-TN5G," vol. 0, p. 21, 2018. [Online]. Available: https://www.itu.int/dms_pub/itu-t/opb/tut/T-TUT-HOME-2018-2-PDF-E.pdf

[66] Y. Nakayama, D. Hisano, T. Kubo, Y. Fukada, J. Terada, and A. Otaka, "Low-latency routing scheme for a fronthaul bridged network," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 1, pp. 14–23, Jan 2018.

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

[68] D. Sabella, P. Rost, Y. Sheng, E. Pateromichelakis, U. Salim, P. Guitton-Ouhamou, M. D. Girolamo, and G. Giuliani, "Ran as a service: Challenges of designing a flexible ran architecture in a cloud-based heterogeneous mobile network," in *2013 Future Network Mobile Summit*, July 2013, pp. 1–8.

[69] A. Garcia-Saavedra, X. Costa-Perez, D. J. Leith, and G. Iosifidis, "Fluidran: Optimized vran/mec orchestration," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 2366–2374.

[70] D. Wubben, P. Rost, J. S. Partelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and impact of cloud computing on 5g signal processing: Flexible centralization through cloud-RAN," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 35–44, 2014.

[71] M. D. Andrade, M. Tornatore, A. Pattavina, A. Hamidian, and K. Grobe, "Cost models for baseband unit (bbu) hotelling: From local to cloud," in *IEEE CloudNet*, Oct 2015, pp. 201–204.

[72] T. Naveh, "Mobile backhaul: fiber vs. microwave: Case study analyzing various backhaul technology strategies," *White Paper from Ceragon*, 2009. [Online]. Available: http://www.winncom.com/images/stories/Ceragon_Mobile_Backhaul_Fiber_Microwave_WP.pdf

[73] M. Paolini, "An analysis of the total cost of ownership of point-to-point, point-to-multipoint, and fibre options," *White paper on crucial economics for mobile data backhaul*, vol. 56, 2012.

[74] W. Briglauer, S. Frübing, and I. Vogelsang, "The impact of alternative public policies on the deployment of new communications infrastructure–a survey," *Review of Network Economics*, vol. 13, no. 3, pp. 227–270, 2014.

[75] H. Gruber, J. Hätönen, and P. Koutroumpis, "Broadband access in the eu: An assessment of future economic benefits," *Telecommunications Policy*, vol. 38, no. 11, pp. 1046–1058, 2014.

[76] A. Garcia-Saavedra, G. Iosifidis, X. Costa-Perez, and D. J. Leith, "Joint optimization of edge computing architectures and radio access networks," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 11, pp. 2433–2443, 2018.

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

[78] N. Carapellese, M. Tornatore, A. Pattavina, and S. Gosselin, "Bbu placement over a wdm aggregation network considering otn and overlay fronthaul transport," in *2015 European Conference on Optical Communication (ECOC)*. IEEE, 2015, pp. 1–3.

[79] C. Ranaweera, E. Wong, A. Nirmalathas, C. Jayasundara, and C. Lim, "5g c-ran with optical fronthaul: An analysis from a deployment perspective," *Journal of Lightwave Technology*, vol. 36, no. 11, pp. 2059–2068, 2017.

[80] V. Suryaprakash and G. P. Fettweis, "An analysis of backhaul costs of radio access networks using stochastic geometry," in *2014 IEEE International Conference on Communications (ICC)*. IEEE, 2014, pp. 1035–1041.

[81] D. Sesto-Castilla, E. Garcia-Villegas, G. Lyberopoulos, and E. Theodoropoulou, "Use of machine learning for energy efficiency in present and future mobile networks," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.

[82] C. V. Nahum, L. D. N. M. Pinto, V. B. Tavares, P. Batista, S. Lins, N. Linder, and A. Klautau, "Testbed for 5g connected artificial intelligence on virtualized networks," *IEEE Access*, vol. 8, pp. 223 202–223 213, 2020.

[83] R. Gupta and S. Kalyanasundaram, "Resource allocation for self-backhauled networks with half-duplex small cells," in *ICC Workshops*, May 2017, pp. 198–204.

[84] R. Jain, D.-M. Chiu, and W. R. Hawe, *A quantitative measure of fairness and discrimination for resource allocation in shared computer system*. Eastern Research Laboratory, Digital Equipment Corporation Hudson, MA, 1984, vol. 38.

[85] S. L. Loyka, "Channel capacity of mimo architecture using the exponential correlation matrix," *IEEE Communications letters*, vol. 5, no. 9, pp. 369–371, 2001.

[86] N. Palizban, S. Szyszkowicz, and H. Yanikomeroglu, "Automation of millimeter wave network planning for outdoor coverage in dense urban areas using wall-mounted base stations," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 206–209, 2017.

[87] G. Salami, O. Durowoju, A. Attar, O. Holland, R. Tafazolli, and H. Aghvami, "A comparison between the centralized and distributed approaches for spectrum management," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 2, pp. 274–290, 2010.

[88] F. Shah-Mohammadi and A. Kwasinski, "Fast learning cognitive radios in underlay dynamic spectrum access: Integration of transfer learning into deep reinforcement learning," in *2020 Wireless Telecommunications Symposium (WTS)*. IEEE, 2020, pp. 1–7.

[89] P. Shetkar and S. B. Ronghe, "Spectrum sensing and dynamic spectrum allocation for cognitive radio network," in *2018 4th International Conference for Convergence in Technology (I2CT)*. IEEE, 2018, pp. 1–5.

[90] W. Zhang, L. Deng, and Y. C. Kiat, "Dynamic spectrum allocation for heterogeneous cognitive radio network," in *2016 IEEE Wireless Communications and Networking Conference*. IEEE, 2016, pp. 1–6.

[91] H. Zhang, N. Yang, W. Huangfu, K. Long, and V. C. Leung, "Power control based on deep reinforcement learning for spectrum sharing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 6, pp. 4209–4219, 2020.

[92] P. Lv, M. Fu, Y. Zhuo, H. Zhao, and J. Zhang, "A dynamic spectrum access method based on q-learning," in *2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)*. IEEE, 2020, pp. 135–141.

[93] K. Sharma, A. Rana, and B. Aneja, "A fuzzy-logic based framework for dynamic channel allocation with improved transmission in cognitive radio," in *2016 International Conference on Signal Processing and Communication (ICSC)*, 2016, pp. 31–36.

[94] H. Cha and S.-L. Kim, "A reinforcement learning approach to dynamic spectrum access in internet-of-things networks," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.

[95] F. Bernardo, R. Agustí, J. Pérez-Romero, and O. Sallent, "Intercell interference management in ofdma networks: A decentralized approach based onreinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 6, pp. 968–976, 2011.

[96] F. Bernardo, R. Agustí, J. Pérez-Romero, and O.Sallent, "An application of reinforcement learning for efficient spectrum usage in next-generation mobile cellular networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 4, pp. 477–484, 2010.

[97] B. Bojović, E. Meshkova, N. Baldo, J. Riihijärvi, and M. Petrova, "Machine learning-based dynamic frequency and bandwidth allocation in self-organized lte dense small cell deployments," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–16, 2016.

[98] Y. A. Abohamra, M. R. Soleymani, and Y. R. Shayan, "Using beamforming for dense frequency reuse in 5g," *IEEE Access*, vol. 7, pp. 9181–9190, 2019.

[99] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, "Access-aware backhaul optimization in 5g," in *Proceedings of the 16th ACM International Symposium on Mobility Management and Wireless Access*. ACM, 2018, pp. 124–127.

[100] Y. Zhang, D. He, W. He, Y. Xu, Y. Guan, and W. Zhang, "Dynamic spectrum allocation by 5g base station," in *2020 International Wireless Communications and Mobile Computing (IWCMC)*, 2020, pp. 1463–1467.

[101] M. Srinivasan, V. J. Kotagi, and C. S. R. Murthy, "A q-learning framework for user qoe enhanced self-organizing spectrally efficient network using a novel inter-operator proximal spectrum sharing," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2887–2901, 2016.

[102] G. Zhang, K. Yang, J. Wei, K. Xu, and P. Liu, "Virtual resource allocation for wireless virtualization networks using market equilibrium theory," in *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2015, pp. 366–371.

[103] T. X. Vu, S. Chatzinotas, V.-D. Nguyen, D. T. Hoang, D. N. Nguyen, M. Di Renzo, and B. Ottersten, "Machine learning-enabled joint antenna selection and precoding design: From offline complexity to online performance," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3710–3722, 2021.

[104] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5438–5453, 2018.

[105] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink mimo," *IEEE Access*, vol. 7, pp. 7599–7605, 2018.

[106] T. E. Bogale, X. Wang, and L. B. Le, "Adaptive channel prediction, beamforming and scheduling design for 5g v2i network: Analytical and machine learning approaches," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 5, pp. 5055–5067, 2020.

[107] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep reinforcement learning for 5g networks: Joint beamforming, power control, and interference coordination," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1581–1592, 2019.

[108] G. Cerar, H. Yetgin, M. Mohorčič, and C. Fortuna, "Machine learning for wireless link quality estimation: A survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 696–728, 2021.

[109] J.-R. Jiang, "Short survey on physical layer authentication by machine-learning for 5g-based internet of things," in *202020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII)*. IEEE, 2020, pp. 41–44.

[110] A. Mughees, M. Tahir, M. A. Sheikh, and A. Ahad, "Towards energy efficient 5g networks using machine learning: Taxonomy, research challenges, and future research directions," *IEEE Access*, vol. 8, pp. 187 498–187 522, 2020.

[111] Y. L. Lee and D. Qin, "A survey on applications of deep reinforcement learning in resource management for 5g heterogeneous networks," in *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2019, pp. 1856–1862.

[112] M. E. M. Cayamcela and W. Lim, "Artificial intelligence in 5g technology: A survey," in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, 2018, pp. 860–865.

[113] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2392–2431, 2017.

[114] S. K. Sharma and X. Wang, "Toward massive machine type communications in ultra-dense cellular iot networks: Current issues and machine learning-assisted solutions," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 426–471, 2019.

[115] Y. Xu, W. Xu, F. Yin, J. Lin, and S. Cui, "High-accuracy wireless traffic prediction: A gp-based machine learning approach," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, pp. 1–6.

[116] U. Paul, L. Ortiz, S. R. Das, G. Fusco, and M. M. Buddhikot, "Learning probabilistic models of cellular network traffic with applications to resource management," in *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, 2014, pp. 82–91.

[117] H. Maciejewski, M. Sztukowski, and B. Chowanski, "Traffic profiling in mobile networks using machine learning techniques," in *International Conference on Digital Information Processing and Communications*. Springer, 2011, pp. 132–139.

[118] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ann) modeling and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 22, no. 5, pp. 717–727, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0731708599002721

[119] T. Liu, Y. Ye, S. Yin, H. Chen, G. Xu, Y. Lu, and Y. Chen, "Digital predistortion linearization with deep neural networks for 5g power amplifiers," in *2019 European Microwave Conference in Central Europe (EuMCE)*, 2019, pp. 216–219.

[120] H. Yang, B. Wang, Q. Yao, A. Yu, and J. Zhang, "Efficient hybrid multi-faults location based on hopfield neural network in 5g coexisting radio and optical wireless networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 4, pp. 1218–1228, 2019.

[121] R. Khdhir, B. Cousin, K. Mnif, and K. Ben Ali, "Neural network approach for component carrier selection in 4g/5g networks," in *2018 Fifth International Conference on Software Defined Systems (SDS)*, 2018, pp. 112–117.

[122] B. Tian, Q. Zhang, X. Xin, Q. Tian, X. Wu, Y. Tao, Y. Shen, G. Cao, and N. Liu, "Recursive neural network based rrh to bbu resource allocation in 5g fronthaul network," in *2018 Asia Communications and Photonics Conference (ACP)*, 2018, pp. 1–3.

[123] X. Zhao, F. Du, S. Geng, N. Sun, Y. Zhang, Z. Fu, and G. Wang, "Neural network and gbsm based time-varying and stochastic channel modeling for 5g millimeter wave communications," *China Communications*, vol. 16, no. 6, pp. 80–90, 2019.

[124] B. Shubyn and T. Maksymyuk, "Intelligent handover management in 5g mobile networks based on recurrent neural networks," in *2019 3rd International Conference on Advanced Information and Communications Technologies (AICT)*, 2019, pp. 348–351.

[125] A. Mazin, M. Elkourdi, and R. D. Gitlin, "Accelerating beam sweeping in mmwave standalone 5g new radios using recurrent neural networks," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, 2018, pp. 1–4.

[126] A. M. Mahmood, A. Al-Yasiri, and O. Y. Alani, "Cognitive neural network delay predictor for high speed mobility in 5g c-ran cellular networks," in *2018 IEEE 5G World Forum (5GWF)*, 2018, pp. 93–98.

[127] Q. Yu, Y. Guan, Y. Yu, and C. Yu, "Genetic algorithm optimized back propagation neural networks in behavioral modeling of power amplifiers excited by 5g nr signal," in *2020 9th Asia-Pacific Conference on Antennas and Propagation (APCAP)*, 2020, pp. 1–2.

[128] A. Waqar, S. Khadim, A. Zeb, S. Amir, and I. Khan, "A survey on cognitive radio network using artificial neural network," *International Journal of Future Generation Communication and Networking*, vol. 10, no. 11, pp. 11–18, 2017.

[129] H. D. Trinh, L. Giupponi, and P. Dini, "Mobile traffic prediction from raw data using lstm networks," in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2018, pp. 1827–1832.

[130] G. Zhang, H. Zhou, C. Wang, H. Xue, J. Wang, and H. Wan, "Forecasting time series albedo using narnet based on eemd decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3544–3557, 2020.

[131] M. Das and S. K. Ghosh, "Data-driven approaches for meteorological time series prediction: a comparative study of the state-of-the-art computational intelligence techniques," *Pattern Recognition Letters*, vol. 105, pp. 155–164, 2018.

[132] R. I. Rony, E. Lopez-Aguilera, and E. Garcia-Villegas, "Cooperative spectrum sharing in 5g access and backhaul networks," in *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2018, pp. 239–246.

[133] ETSI, "TS: 138 104 V15.2.0 - 5G; NR; Base Station (BS) radio transmission and reception," *3GPP TS 38.104 version 15.2.0 Release 15*, no. January, p. 10, 2018. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/138100_138199/138104/15.02.00_60/ts_138104v150200p.pdf

[134] O. Liberg, M. Sundberg, Y.-P. E. Wang, J. Bergman, and J. Sachs, "Chapter 5 - lte-m," in *Cellular Internet of Things*, O. Liberg, M. Sundberg, Y.-P. E. Wang, J. Bergman, and J. Sachs, Eds. Academic Press, 2018, pp. 135–197. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128124581000058

[135] P. A. Adedeji, S. Akinlabi, O. Ajayi, and N. Madushele, "Non-linear autoregressive neural network (narnet) with ssa filtering for a university energy consumption forecast," *Procedia Manufacturing*, vol. 33, pp. 176–183, 2019.

[136] N. Salhab, R. Rahim, R. Langar, and R. Boutaba, "Deep neural networks approach for power head-room predictions in 5g networks and beyond," in *2020 IFIP Networking Conference (Networking)*. IEEE, 2020, pp. 579–583.

[137] G. Zhang, H. Zhou, C. Wang, H. Xue, J. Wang, and H. Wan, "Forecasting time series albedo using narnet based on eemd decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3544–3557, 2020.

[138] B. Cha and S.-K. Noh, "Learning using lte rsrp and narnet in the same indoor area," in *2019 23rd International Computer Science and Engineering Conference (ICSEC)*. IEEE, pp. 261–264.

[139] A. He and X. Jin, "Narnet-based prognostics modeling for deteriorating systems under dynamic operating conditions," in *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2018, pp. 1322–1327.

[140] M. Kayri, "Predictive abilities of bayesian regularization and levenberg–marquardt algorithms in artificial neural networks: a comparative empirical study on social data," *Mathematical and Computational Applications*, vol. 21, no. 2, p. 20, 2016.

[141] N. Docomo, "Docomo 5g white paper," *White Paper,(Jul. 2014)*, 2014.

[142] A. Haidine, F. Z. Salmam, A. Aqqal, and A. Dahbi, "Artificial intelligence and machine learning in 5g and beyond: a survey and perspectives," *Moving Broadband Mobile Communications Forward: Intelligent Technologies for 5G and Beyond*, p. 47, 2021.