Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

PhD Thesis

Massive MIMO and Millimetre Wave Technologies: Design, Application and Integration with ML Techniques for 5G and Beyond Networks

Author:
Andrea Bonfante

Supervisors:
Prof. Nicola Marchetti
Dr. Lorenzo Galati Giordano

6th July 2021

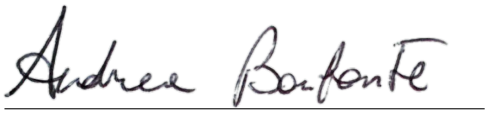
Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Signed:



Andrea Bonfante, 6th July 2021

Abstract

Massive Multiple-Input-Multiple-Output (mMIMO) and millimetre Wave (mmWave) are key enabling technologies to realise omnipresent, scalable and dynamic 5-th Generation (5G) and beyond networks that provide seamless wireless connectivity supporting enhanced data rates. Fundamental challenges for the full-scale deployments of these technologies are the possibility to integrate different services, e.g. wireless access and backhaul, and the support of applications like industrial robotics demanding a reliable wireless network service. Significant enablers for rapid adoption are the utilisation of self-Backhauling (s-BH) technologies to integrate access and backhaul services in the same spectrum resources and the adoption of Machine Learning (ML) techniques to add predictive features to the control operation of the network.

In this dissertation, we investigate a mMIMO-based s-BH network architectures to facilitate cost-effective network deployments of Small Cells (SCs), and we propose a ML-based beam recovery method to mitigate dynamic blockages and provide reliable wireless connections in mmWave indoor networks. Our work focuses on addressing the following research questions:

- What is the potential of using mMIMO-based s-BH network architectures to provide backhaul links to SCs and favour Ultra Dense Network (UDN) deployments?
- How to improve the mmWave network data rate stability by making the mmWave links more robust to blockage events?
- How to design novel radio control methods by integrating mmWave systems with ML models to provide beam state predictions?

We adopt several simulative and analytical tools, such as system-level simulations and stochastic geometry analysis, that evaluate large-scale networks performance. Moreover, we employ ML models such as Deep Neural Network (DNN) to make predictions about the

network environment.

First, we investigate through comprehensive Third Generation Partnership Project (3GPP)-based system-level simulations and analytical formulations the User Equipments (UEs) data rate performances of the mMIMO-based s-BH network architecture and compare these results to the mMIMO Direct Access (DA) network architecture, in which mMIMO-Base Stations (BSs) directly serve UEs in the absence of SCs. In this work, we propose a novel deployment of SCs, namely ad-hoc deployment, where SCs are positioned close to UEs to achieve Line-of-Sight (LoS) access links. Then, we evaluate the optimal partition between backhaul and access resources that maximises the end-to-end UE data rates. We show that ultra-dense SCs deployments supported by mMIMO s-BH provide rate improvements for cell-edge UEs that amount to 30% and a tenfold gain compared to mMIMO DA solutions with pilot reuse scheme 3 and reuse scheme 1, respectively. On the other hand, mMIMO s-BH underperforms mMIMO DA above the median value of the UE data rates when the effect of pilot contamination is less severe and DA links achieve a higher LoS probability.

Next, motivated by the need to improve the Sub-6 GHz links capacity further and to provide more stable wireless links, we moved our focus on the blockage problem in mmWave indoor networks, considering a DA network architecture. In this scenario, the mmWave links experience rapid and temporary fluctuations of the received signal power when they encounter blocking objects, such as humans and robots moving in the environment. We propose a novel method that performs the beam switching in advance based on the predictions provided by beam-specific DNN models and avoids the delay introduced by the method based on a detection threshold. We show that during the blocked intervals, the prediction-based method guarantees higher signal level stability and up to 82% data rate improvement to the detection-based method when blockers move at a speed of 2 m/s. In these conditions, the higher frequency of beam switching penalises the detection-based method's data rate while the prediction-based method achieves better results due to the ability of generalise well with different blocker speeds.

Finally, we show an example of a ML workflow for implementing the prediction-based method in the Open-Radio Access Network (O-RAN) architecture. This is because Software (SW)-based components and the use of a ML workflow can be the two main directions to facilitate tuning and optimisation of network parameters and the ML models training.

Acknowledgements

This work would not be possible without the help that I have received from many talented and brilliant people truly passionate about research that I have been lucky to meet during my PhD studies.

I express a profound sense of gratitude to Prof. Nicola Marchetti and Dr Lorenzo Galati Giordano for their valuable guidance, support, constructive insights and the numerous opportunities they provided me during these years.

I express my gratitude to Dr Irene Macaluso, who advised me for the ML chapter, taking the time to meet and discuss despite all the extraordinary events and circumstances. I am also grateful to Dr David López-Pérez, Dr Giovanni Geraci and Dr Adrian Garcia-Rodriguez for their guidance and thoughtful help during the first part of the PhD program.

I would like to thank the examiners Prof. Lei Zhang and Prof. Marco Ruffini for their insightful suggestions, which helped me to improve the presentation of this thesis.

I would like to thank all Professors, Managers, Members of Technical Staff, Senior Researchers and Administrative Staff in CONNECT Centre and Nokia Bell Labs Dublin for being great models to follow and, in general, for having helped and supported me.

Sincere thanks to my colleagues and ex-colleagues in CONNECT Centre and Bell Labs to have spent quality time together and created a lasting legacy of great memories. Thanks, Fadhil, Ramy, Nima, Alan, Jernej, Merim, Erika, Joao, Conor, Boris, Eamonn, Parna, Jean-Baptiste, Andre', Francesco, Eloise, Stefan, Davi, Pieter and many, many more. Thanks to my Italian friend Gianluca, great companion of cycling and trail activities.

I would like to thank my parents Nadia and Maurizio, and my siblings Anna and Marco, who provided me with the comforting and support to wrap-up this thesis on schedule during this exceptionally challenging time. I dedicate this thesis to you and to my beloved girlfriend, Valeria, who first encouraged me to pursue a PhD degree and since then has patiently

supported and motivated me to complete this path.

Finally, I would like to acknowledge the founding received from the Irish Research Council and Nokia Bell Labs Ireland that under the Enterprise Partnership Scheme Grant Number EPSPG/2016/106 provided me the financial resources to conduct my research activities and present my work abroad during my PhD program.

Table of Contents

List of Figures

List of Tables

List of Acronyms

3D	Three-Dimensional
3GPP	Third Generation Partnership Project
4G	4-th Generation
5G	5-th Generation
Adam	Adaptive moment estimation
ADC	Analog-to-Digital converter
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
AR	Augmented Reality
BS	Base Station
CDF	Cumulative Distribution Function
CN	Client Node
C-RAN	cloud RAN
CSI	Channel State Information
DA	Direct Access
DAC	Digital-to-Analog converter
DL	Downlink
DN	Distribution Node
DNN	Deep Neural Network

DoA Direction of Arrival

DoF Degrees-of-Freedom

eMBB Enhanced Mobile BroadBand

FD Full Duplex

FDD Frequency Division Duplex

FDMA Frequency-division multiple access

FN false negative

FP false positive

F-OFDM Filtered Orthogonal Frequency Division Multiplexing

GT Ground Truth

HD Half-Duplex

HetNet Heterogeneous Network

HW Hardware

IAB Integrated Access and Backhaul

IEEE Institute of Electrical and Electronics Engineers

IIoT Industrial Internet of Things

IoT Internet of Things

ISD Inter Site Distance

LiDAR Light Detection And Ranging

LNA Low Noise Amplifier

LoS Line-of-Sight

LSAS Large Scale Antenna System

LTE Long Term Evolution

MBB Mobile BroadBand

MCS Modulation and Coding Scheme

MIMO multiple-input-multiple-output

ML Machine Learning

mMIMO massive Multiple-Input-Multiple-Output

mMIMO-BS mMIMO base station

mMTC massive Machine-Type Communication

mmWave millimetre Wave

mmWave-BS mmWave Base Station

MNO Mobile Network Operator

MPC Multipath Components

NLoS Non Line-of-Sight

NOMA Non-Orthogonal Multiple Access

NR New Radio

NSA Non-Standalone

OFDM Orthogonal Frequency Division Multiplexing

O-RAN Open-Radio Access Network

PA Power Amplifier

PHY Physical layer

RAN Radio Access Network

RB Resource Block

ReLu Rectified Linear unit

RF Radio Frequency

RN Relay Node

RRH Remote Radio Head

RSRP Reference Signal Received Power

RX Receiver

SA Standalone

s-BH self-Backhauling

SC Small Cell

SCM Spatial Channel Model

SINR signal-to-interference-and-noise ratio

SIR signal-to-interference ratio

SNR signal-to-noise ratio

SSB synchronisation signal block

SW Software

TDD Time Division Duplex

TDMA Time-division multiple access

TN true negative

TP true positive

TTI Transmission Time Interval

TX Transmitter

UAV Unmanned Aerial Vehicle

UDN Ultra Dense Network

UE User Equipment

UL Uplink

ULA Uniform Linear Array

UPA Uniform Planar Array

URLLC Ultra-Reliable Low-Latency Communication

VCO Voltage Controlled Oscillator

VR Virtual Reality

Wi-Fi Wireless-Fidelity

1 Introduction

Back in 2012, the telecommunication industry's key stakeholders set the ambitious goal to increase the network's capacity by a factor of 1000× to address the exponential growth of the mobile data traffic that in the early 2010s nearly doubled every year [1]. Both the industry and the research community have responded to this call promoting all the existing paradigms to improve the network capacity by increasing cell density, improving cell spectral efficiency and adopting higher channel bandwidths. A plethora of techniques can tackle these fundamental problems. Among the various enabling technologies for 5-th Generation (5G), in this thesis, we focus on the ultra-densification of the network, massive Multiple-Input-Multiple-Output (mMIMO) and millimetre Wave (mmWave) communication. Ultra Dense Network (UDN) deployments make use of many Small Cells (SCs), which consist of low-power Base Stations (BSs) with reduced coverage area that are used to increase the network capacity [2]. Sub-6 GHz mMIMO increases the number of BS antennas to enable the simultaneous transmission to multiple User Equipments (UEs) spatially multiplexed in the same time-frequency resources [3]. Lastly, mmWave communication can significantly increase the transmission bandwidth, utilising the spectrum available at mmWave frequencies [4].

The 5G network's roll-out is currently advancing, and many 5G network deployments are already operational and focus on enhancing the Mobile BroadBand (MBB) service [5]. Nevertheless, the 5G technology standard covers several other requirements besides scaling the network capacity and improving the network data rate [6]. One of the main challenges of network densification is represented by the high investments required for the BS deployment, especially for installing the fibre optics cable connecting the core network to the Radio Access Network (RAN), representing the higher portion of the BS installation's costs [7]. Therefore, other functionalities of 5G networks, such as the option to use self-

Backhauling (s-BH), also known as Integrated Access and Backhaul (IAB), are included in the 5G standard [8]. Moreover, many stakeholders in the industry demand low latency and high-reliability communications and look forward to using these new 5G connections for private networks with applications to Industrial Internet of Things (IIoT), network control system and mission-critical communications [9]. The high spectrum available at mmWave frequencies can achieve gigabit per second connections [4]. On the other hand, given the sensitivity of mmWave communication to blockage, there are several interrogations about whether the mmWave can support stable and reliable connections [10]. New approaches based on Machine Learning (ML) tools shown to be effective for predicting blockage events [11, 12]. However, there is a lack of potential solutions that integrates these methods into the mmWave network operations.

All in all, a key aspect of the 5G wireless communication network is integrating different services and technologies to provide seamless connectivity. For instance, the 5G technology s-BH integrates access and backhaul services to provide a cost-effective option to enable massive and flexible deployments of SCs. At the same time, the use of ML tools in wireless communication is envisioned as an essential step towards optimising 5G networks and designing future RANs [13]. ML models can be employed in a wireless network to manage the control operations proactively to improve throughput and reduce latency [14]. The key challenge is combining various 5G technologies, e.g. s-BH and mMIMO, while integrating new enabling technologies, e.g. ML, to evolve the existing 5G network infrastructure to be more flexible, massively deployable and proactively so that it can boost the 5G network data rate while providing a reliable wireless network service.

1.1 Scope

In this dissertation, we investigate novel network architectures and methods resulting from the crossover between three technologies for 5G (mMIMO, UDN and mmWave) and other emerging techniques like s-BH and ML. The general research question that we want to address can be stated as follows: **"How will the combination of 5G technologies and other emerging techniques impact the design of novel network architectures and functionalities to deliver seamless wireless connectivity that supports multiple**

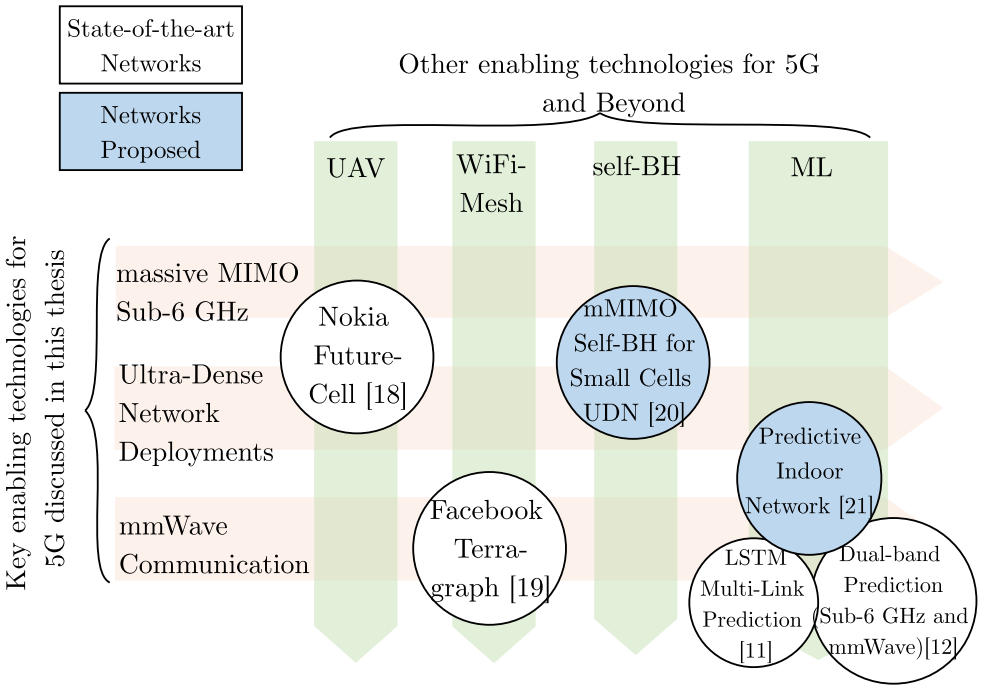


Figure 1.1: Examples of state-of-the-art networks and methods resulting from the crossover between various enabling technologies for 5G and other enabling technologies for 5G and beyond. The works proposed in this dissertation are represented with a darker background colour.

5G services in the years to come?" Our discussion will touch on the various technologies of 5G: mMIMO, UDN and mmWave, as considered in [15]. We investigate mMIMO for s-BH to facilitate UDN deployments of SCs, and we propose a ML-based method for beam recovery in mmWave communication to improve the data rate stability of the mmWave links.

In order to better visualise the proposed networks and methods, we represented in the rows of Fig. 1.1 the three 5G technologies: mMIMO, UDN and mmWave. At the same time, we represented in the columns of Fig. 1.1 other technologies such as Unmanned Aerial Vehicle (UAV), Wireless-Fidelity (Wi-Fi)-Mesh, s-BH and ML that are becoming increasingly important for evolving 5G wireless communication networks [14, 16, 17]. In the cross cells between rows and columns, we represent different examples of state-of-the-art networks and methods that combine 5G technologies with other enabling technologies.

As shown in Fig. 1.1, Nokia F-Cell considers technologies such mMIMO and UAV, which are used to facilitate UDN deployment of SCs through mMIMO wireless backhaul and UAV-

based delivery [18]. Facebook Terragraph integrates Wi-Fi-Mesh technology into mmWave Base Stations (mmWave-BSs) to reduce the infrastructure costs of mmWave networks roll-out [19]. Additionally, the methods introduced in [11, 12] apply ML-based tools to predict the blockage events in mmWave networks.

At the same time, the networks and methods proposed in this dissertation are depicted in Fig. 1.1 with a darker background colour. In [20], we propose to use mMIMO and s-BH technologies together for UDN deployments of SCs. Then, in [21], we integrate ML models in mmWave-BSs to propose a predictive approach to recover the beam in an indoor mmWave network. Combining 5G technologies and integrating other enabling technologies is challenging to be realised in practice, and several problems related to the use of the 5G technologies arise. Therefore, throughout the dissertation, our focus will be on answering the following three research questions linked to the general research problem:

Q-1: What are the end-to-end UEs data rates performance when adopting the mMIMO-based s-BH network architecture? When to use mMIMO for s-BH? or when to use mMIMO to provide Direct Access (DA) service to UEs?

Due to the scarcity of spectrum at Sub-6 GHz bands, in-band s-BH operations aim to integrate access and backhaul services to favour wide-scale deployments of SCs with wireless backhaul connections. In Chapter 3, we employ mMIMO for s-BH to serve multiple SCs in parallel. Our scope is to study the most favourable strategy for deploying self-backhauled SCs in the network, learn how to manage the backhaul and access operations carefully and compare two mMIMO network architectures: 1) to provide wireless in-band backhaul for SCs acting as Relay Nodes (RNs) to UEs, and 2) to provide DA to UEs.

Q-2: What are the effects of dynamic blockers when using mmWave for DA communications in an indoor scenario? How to improve the throughput stability by making the mmWave links more robust to blockage events?

One of the main challenges for mmWave networks is the sensitivity of mmWave links to blockage. We address these questions in Chapter 4, where the scope is to gain insights about the blockage effect's temporal dynamics when mobile blockers, moving around in an

indoor environment, cause frequent obstructions of the mmWave links. Additionally, we employ a beam recovery method based on detection reacting to the blockage events and switching to a backup direction. Our scope in Chapter 4 is to study the mmWave network's data rate and latency performance, evaluating the blockage mitigation capabilities of the detection approach.

Q-3: How to avoid the delay for switching to the backup beam at every blockage event in indoor mmWave DA communications? How to design novel radio control methods by deploying ML models into the mmWave networks for making beam state predictions?

Our scope in Chapter 5 is to adopt ML algorithms to make predictions about the blockage state of the mmWave beams and use these predictions in a control mechanism that switches the beam proactively. The main challenge is to avoid the delay introduced by conventional reactive methods by leveraging on the predictions provided by a ML tools. Overall, in Chapter 5, we assess the feasibility to deploy ML models into the mmWave network operations by considering a realistic multi-UE scenario of mmWave network deployment.

1.2 Structure

The remaining chapters of this dissertation are organised as follows. In Chapter 2, we provide background concepts to explain the various 5G technologies discussed in this thesis and present the state-of-the-art networks and methods. In Chapter 3, we analyse a mMIMO-based s-BH network architecture for UDN deployments. In Chapter 4, we study the mmWave DA indoor network performance under dynamic blockage effect by employing a detection method. In Chapter 5, we propose a technique that integrates ML tools into a mmWave indoor network to predict the blockage events and apply preemptive beam switching. Finally, in Chapter 6, we summarise the main findings discussed in this dissertation, and we show several future directions to further develop the 5G networks.

Chapter 2 – Background on 5G Technologies, State-of-art Networks and Methods

In Chapter 2, we introduce various radio technologies employed in 5G network architectures: mMIMO, UDN and mmWave communication. We discuss the main characteristics of these three technologies and their main drawbacks. For simplifying the network rollout, state-of-the-art networks integrate new enabling technologies such as s-BH, UAV and Wi-Fi-Mesh. On the other hand, novel approaches to be applied in mmWave networks consider the real-time detection of blockages or the utilisation of ML tools to predict these events. Throughout the chapter, we provide a detailed description of each state-of-the-art network and method. We conclude the chapter by giving some final remarks, where we highlight the connection between the background works and the remaining chapters of this dissertation.

Chapter 3 – Massive MIMO Self-Backhauling for Ultra-Dense Small Cell Deployments

Self-BH, recently included in the Third Generation Partnership Project (3GPP) New Radio (NR) Standard, consists of an integrated access and backhaul network architecture that aims to reduce the network roll-out investments required to extend the wired backhaul infrastructure and acquire additional spectrum dedicated to the wireless backhaul. In Chapter 3, we use mMIMO to provide in-band wireless backhauling to a dense deployment s-BH SCs acting as communication RNs to UEs. We show the achievable end-to-end UE rates with this architecture by adopting different SCs deployment strategies. We finally compare the mMIMO-based s-BH network architecture to mMIMO DA ones, where mMIMO provides wireless DA to UEs.

Chapter 4 – Analysis of mmWave Indoor Network Performance with Detection of Dynamic Blockers

In Chapter 4, we study an indoor mmWave cellular network, and we analyse the access links' performance when dynamic blockers moving in the environment temporally obstruct the beam pair established with the UE. We consider that each mmWave-BS employs a 2D

beamforming codebook containing multiple beamforming vectors covering different parts of the mmWave-BS sector area. To contrast the significant data rate losses occurring during the blocked time-instants, we introduce a radio link control method, which detects the blockage events when the received signal power drops below a given threshold. At this point, the detection triggers the beam recovery mechanism that searches for a non-blocked beam to re-configure the transmission.

Chapter 5 – Predictive Indoor mmWave Networks with Dynamic Blockage

In Chapter 5, we analyse the integration of ML tools into a mmWave indoor network control operations of the access service that copes with the switching delay introduced by reactive methods. We propose a novel prediction-based method that utilises Deep Neural Network (DNN) models to proactively manage the switching operation of mmWave networks. This novel approach is evaluated in a multi-UE scenario, where UEs are assigned to multiple beams of the codebook, and it is compared to a beam detection method, which detects a blocked beam based on a detection threshold.

Chapter 6 – Conclusion and Future Works

This last chapter summarises the main contributions of the dissertation. We provide the concluding remarks that tie the key takeaways together and enlighten the main trends of network design for the years ahead.

1.3 Contributions

Chapter 2 – Background on 5G Technologies, State-of-art Networks and Methods

Our contribution in Chapter 2 can be summarised as follows:

- We present various technologies forming the 5G networks: mMIMO, UDN, mmWave communication and s-BH. Then, we explain the benefits and drawbacks of employing each technology in the 5G networks.

- We introduce several examples of state-of-the-art networks and methods related to the network architecture and the control method proposed in this dissertation. Afterward, we highlight how the technical contributions of this dissertation progress beyond state-of-the-art networks and methods.

Chapter 3 – Massive MIMO Self-Backhauling for Ultra-Dense Small Cell Deployments

The contributions of the chapter are summarised as follows:

- We propose a mMIMO-based s-BH architecture with an ad-hoc SCs deployment and we investigate the UE data rate performance changing the distance from the SC to the UE and the SC antenna pattern. Then, we analyse the same architecture with a random SCs deployment, varying the number of SCs in the network.
- We study the effect of the backhaul/access partition variation in a mMIMO-based s-BH architecture and investigate the maximum achievable end-to-end UE data rates while optimising the backhaul/access resource allocation.
- We compare the mMIMO s-BH and DA network architectures performance, accounting for the pilot contamination effect. We also analyse the UEs performance at different throughput percentiles to study the best network architectures for cell-edge UEs.

Chapter 4 – Analysis of mmWave Indoor Network Performance with Detection of Dynamic Blockers

In Chapter 4, our contributions are summarised as follows:

- We analyse the mmWave DA communication in an indoor network scenario affected by dynamic blockages through system-level simulations. We adopt a geometric blockage model to gain more insights about the stability of the UEs data rate under the temporal transition from Line-of-Sight (LoS) to Non Line-of-Sight (NLoS) channel propagation conditions due to the blockage effect.

- We introduce a beam recovery method in the mmWave network operations that performs real-time detection of blockage events and recovers the mmWave connections by switching to an alternative beam in case of blockage.
- We provide a detailed analysis of the UEs performance testing key aspects that characterise the dynamic blockage effect, such as the blocker speed and the blockage probability.

Chapter 5 – Predictive Indoor mmWave Networks with Dynamic Blockage

Our contributions in Chapter 5 can be summarised as follows:

- We propose a novel multi-UE prediction-based method for beam recovery that incorporates the beam measurements from multiple BS to predict the subsequent blockage events. This early indication enables to initiate the beam switching in advance and complete this operation by the time the mmWave link becomes blocked.
- We validate the prediction-based method by deploying the DNN models online into a 3GPP-based system-level simulator to verify the close match to the performance of an ideal method that has perfect knowledge of the future beam states. We then compare the prediction method to other two cases: *i)* adopting the beam recovery method based on detection and *ii)* utilising a fixed beam method. Finally, we show that the prediction method switches to a backup beam earlier than the blockage event, and avoids the data rate loss that occurs with the detection method due to the delay in switching to the backup beam and without the switching for the fixed beam method.
- We provide a quantitative analysis of the three methods' data rate performance with different blocker speeds.

1.4 Dissemination

The papers relevant for this dissertation are listed as follows:

- A. Bonfante, L. Galati Giordano, D. López-Pérez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, M. M. Butt, M. Dzaferagic and N. Marchetti, "Performance of massive

MIMO self-backhauling for ultra-dense small cell deployments", *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1-7.

- A. Bonfante, L. Galati Giordano, D. López-Pérez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, M. M. Butt, and N. Marchetti, "5G massive MIMO architectures: Self-backhauled small cells versus direct access", *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10003-10017, 2019.
- A. Bonfante, L. Galati Giordano, I. Macaluso, and N. Marchetti, "Performance of predictive indoor mmWave networks with dynamic blockers", submitted to *IEEE Transactions on Cognitive Communications and Networking*, Apr. 2021. [Online]. Available: http://arxiv.org/abs/2104.04623v1

The following paper is the outcome of the collaboration with other researchers, and it is not part of the dissertation:

- F. Firyaguna, A. Bonfante, J. Kibiłda, N. Marchetti, "Performance Evaluation of Scheduling in 5G-mmWave Networks under Human Blockage", to be submitted to *2021 IEEE Global Communications Conference Workshops (GLOBECOM Workshops)*.

2 Background on 5G Technologies, State-of-art Networks and Methods

In this chapter, we first present the key directions to improve the network capacity, and we introduce various radio technologies for 5-th Generation (5G) networks: massive Multiple-Input-Multiple-Output (mMIMO), Ultra Dense Network (UDN), millimetre Wave (mmWave) communication and self-Backhauling (s-BH) (also known as Integrated Access and Backhaul (IAB)). We discuss the main differentiators that characterise the state-of-the-art networks and distinguish them from the past 4-th Generation (4G) networks requirements, which were mainly focused on high-speed data rate communications. Then, we show several examples of state-of-the-art networks that combine 5G technologies with new enabling technologies such as Unmanned Aerial Vehicle (UAV) and Machine Learning (ML). For instance, Nokia F-Cell aims to facilitate UDN deployment through wireless backhauled Small Cells (SCs) installed by an UAV [18]. Facebook Terragraph deploys mmWave Base Stations (mmWave-BSs) with Wireless-Fidelity (Wi-Fi)-Mesh technology to facilitate mmWave networks roll-out and reduce the backhaul infrastructure costs [19]. Finally, we move our focus on mmWave Direct Access (DA) communications and we present state-of-the-art beam recovery methods based on detection and blockage prediction approaches based on ML. We provide a detailed description of each network and method, and we explain those particular aspects that make our works about mMIMO for s-BH network architectures and the ML method for beam recovery in mmWave indoor networks to evolve beyond state-of-the-art networks and methods.

2.1 Overview of the Various 5G Technologies for State-of-the-art Networks

The area throughput, which measures the capacity of the network in the Base Station (BS) cell area in bps/km^2, can be expressed as $R_{\text{Cell}} = \rho_b \times \text{SE} \times BW$ [22]. It combines a mixture of the three different factors: the density of BSs denoted as ρ_b and measured in Cells/km^2, the cell spectral efficiency denoted as SE and measured in bps/Hz/Cell and, the channel bandwidth denoted as BW and measured in Hz.

Therefore to improve the area throughput for a network, there are three main directions: increasing the average cell density, improving the cell spectral efficiency and adopting a larger channel bandwidth [22]. Different technologies can be applied to pursue each of these directions. We provide several examples as follows:

- To increase the cell density, Heterogeneous Networks (HetNets) consider two tiers networks formed by a layer of macro BSs providing extended coverage and overlying a layer of SCs, which are BSs with low radiated power that is usually kept in the range 250 mW – 1 W [23]. When the number of SCs per square km deployed in the network becomes very high, the network densification process leads to ultra-dense SC deployments. These have a high density of SCs deployed without network planning and require dedicated spectrum resources to avoid interference from the macro BSs [2].
- To improve the cell spectral efficiency, BSs can transmit with coordinated multi-point transmission/reception techniques [24] and dynamic frame structures [25]. BSs can also employ advanced radio access techniques, e.g. Non-Orthogonal Multiple Access (NOMA) [26], and waveforms, e.g. Filtered Orthogonal Frequency Division Multiplexing (F-OFDM) [27], or be equipped with a mMIMO technology [28].
- To increase the channel bandwidth, the Full Duplex (FD) technology allows Uplink (UL)-Downlink (DL) bidirectional transmission and doubles the available spectrum resources [29]. Moreover, given the scarce availability of spectrum in the Sub-6 GHz bands, one promising direction is to use higher frequency bands such as mmWave frequencies [10].

Among all the techniques mentioned above, in this dissertation, we discuss UDN, mMIMO and mmWave technologies as they are considered fundamental for 5G [15]. We provide a detailed description of each of the three 5G technologies in the following sections. Additionally, we introduce the s-BH solution as part of the 5G technologies considered in this dissertation.

2.1.1 Ultra-dense Small Cell Deployments

Since the first generation of wireless communication systems, Mobile Network Operators (MNOs) have progressively increased the number of BSs in the network to improve capacity. Hence the wireless network deployments have evolved towards considering HetNets architectures formed by a macro BSs layer overlying a layer of SCs, which transmit with less power than the macro BSs and have smaller cell radii [23]. Over the years, the number of 4G SCs sold has exceeded the number of 4G macro BSs. Today SCs represent the primary solution for network densification. The SCs will become even more necessary to ensure coverage for 5G and beyond mmWave network deployments. Indeed, for the higher pathloss affecting these frequency bands, the coverage can be guaranteed with small Inter Site Distance (ISD) between different BSs. Thus, using mmWave for 5G and beyond networks will implicitly consider UDN [4].

In a nutshell, the fundamental property of network densification is that as the network density increases, the User Equipment (UE) benefits from the proximity with the SC, and the Line-of-Sight (LoS) probability improves while the pathloss decreases [30]. This, in turns, provides significant enhancements to the received signal power, and the signal-to-interference-and-noise ratio (SINR) improves. More importantly, each SC has to serve a reduced number of UEs in the channel bandwidth. Because of the improved SINR and mainly because more spectrum resources are available for each UE, network densification provides substantial improvements to the UE throughput, and there is an optimum number of BS which maximise the area spectral efficiency [31].

Nevertheless, several challenges, such as backhaul infrastructure, UE mobility, and SCs deployment planning, have penalised the spreading of ultra-dense deployments of SCs [2]. More specifically, according to the MNOs, the cost-effective availability for backhaul is

currently the main limitation holding back outdoor SCs deployment [7]. Wireless backhaul technologies represent a flexible and cost-efficient solution since they allow deploying an SC without the need for a wired infrastructure for each deployment site [32, 33]. Thus, it opens the possibility to move the SC closer to a cluster of UEs that require a high throughput connection [34].

2.1.2 Massive MIMO Sub-6 GHz

The mMIMO technology was initially proposed by Prof. Thomas Marzetta in 2010 while working at Bell Labs and considers equipping a BS with a large number of antennas and signal processing functions integrated into a Large Scale Antenna System (LSAS). The LSAS enables to simultaneously use of the same set of time-frequency resources that are shared between different devices to increase the spectrum efficiency massively [35]. Since 2010, significant research advancements have been conducted, firstly, to study this technology theoretically [3], then to investigate its applications such as for backhaul and in the unlicensed band [36–40]. At the same time, mMIMO shown its feasibility and practical potential with the first proof-of-concept [41] and, more recently, this technology has been transformed into commercial products [42].

We focus on applying mMIMO for Sub-6 GHz bands, where the use of fully digital architectures with one Radio Frequency (RF) chain for each antenna provides the maximum spatial Degrees-of-Freedom (DoF) for spatial multiplexing. The precoding operation in mMIMO systems, which consists of mapping the multiple transmission data streams into the set of RF processing chains, can adopt linear processing techniques, e.g., conjugate beamforming and zero-forcing. We provided practical insights about this topic in [43], which analysed the essential baseband components of mMIMO through a Software (SW) and Hardware (HW) co-operated platform. Nevertheless, to calculate the precoder coefficients, the system necessitates knowing the Channel State Information (CSI) at the BS transmitter. For instance, in Time Division Duplex (TDD) systems, where the mMIMO channel reciprocity between UL and DL is usually assumed [3], the CSI is acquired at the BS with a sequence of pilot symbols having a length that grows proportional to the number of UEs. In contrast, in Frequency Division Duplex (FDD) systems, the pilot sequence's

Table 2.1: MmWave bands and industrial standards for mmWave communication.

Carrier frequency (Licensing)	**Industrial Standard**	**Bandwidth**	**Applications**
28 GHz and 39 GHz (Licensed)	3GPP NR [46]	Up to 400 MHz	eMBB URLLC
V-band: 60 GHz (Unlicensed)	IEEE 802.11ay [47]	Up to 2.16 GHz	eMBB
E-band: 71-76 GHz and 81-86 GHz	TBD	250 MHz to 4.5 GHz [48]	TBD

length increases proportionally with the product between the number of BS antennas and UEs, making the training overhead larger for FDD than TDD [44].

One of the main problems of the mMIMO technology in multi-cell networks is pilot contamination [3, 35]. This effect occurs when the CSI estimation is contaminated by the other UEs transmitting in different cells and using the same set of orthogonal pilot sequences. Given the high impact of the pilot contamination on the inter-cell interference in a multi-cell scenario, it is important to account for this effect when modelling the mMIMO system performance. In Chapter 4, we account for the pilot contamination effect and we introduce the pilot reuse schemes adopted in [45] to mitigate pilot contamination.

2.1.3 MmWave Communication

By definition, the mmWave bands cover the range of frequencies between 30 and 300 GHz; however, by convention, several industrial standards extend this range to frequencies above 24 GHz. It is worth mentioning that the mmWave bands reserved for mobile communication are licensed for 28 GHz and 39 GHz bands, the 60 GHz band (V-band) is allocated for unlicensed use, and the licensing for the bands from 71-76 GHz and 81-86 GHz (E-band) has to be defined. The Institute of Electrical and Electronics Engineers (IEEE) considered the use of the mmWave band at 60 GHz starting from 2005 when it proposed the IEEE 802.15.3c standard for wireless personal area networks (WPANs) [49]. Then in 2012, the IEEE released the IEEE 802.11ad standard (WiGig), which uses the 60 GHz unlicensed band for serving wireless local area networks (WLANs) [50]. Recently, the IEEE 802.11ad standard has evolved to IEEE 802.11ay, which supports an even higher data rate thanks to the wider channel bandwidth and the multiple-input-multiple-output (MIMO) technique

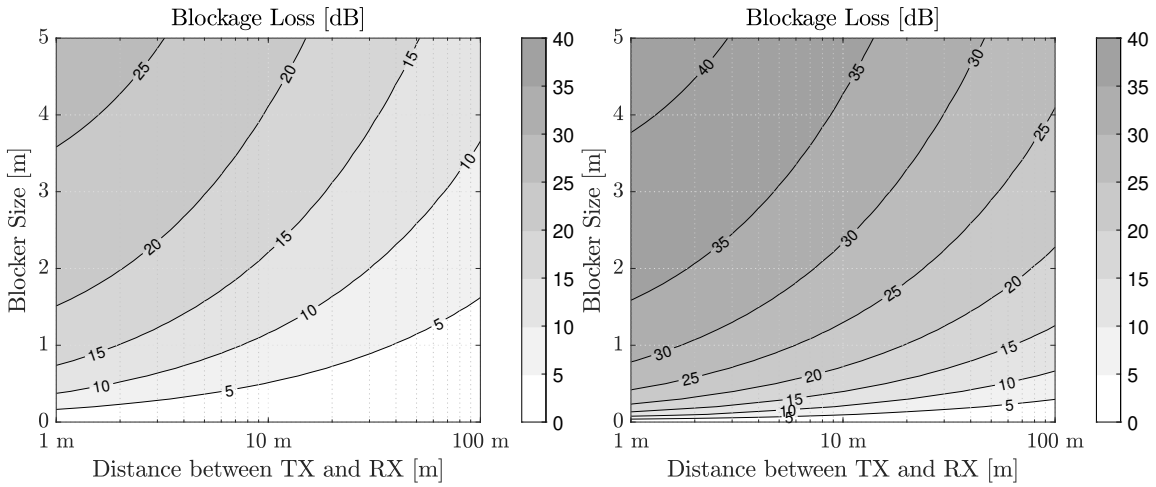
[47].

At the same time, 5G cellular communication systems that usually adopted Sub-6 GHz bands favouring radio propagation conditions started recently using mmWave bands. The main driver has been the availability of large channel bandwidths in the spectrum above 24 GHz. This has pushed for an increase in R&D activities around RF systems operating at mmWave frequencies. Consequently, the Third Generation Partnership Project (3GPP) New Radio (NR) Standard for cellular networks extends the operations from a frequency range 1 (FR1) – below Sub-6 GHz bands – to a second frequency range (FR2), which includes the 28 GHz and 39 GHz mmWave bands [46]. 3GPP NR Standard considers two architectures for FR2 network operations: Non-Standalone (NSA) and Standalone (SA). In the NSA architecture, the mmWave band is used in the user data plane for enhancing the data rate in the area with more traffic demands, while the Sub-6 GHz band is adopted for control signalling. Differently, the SA architecture does not need to rely on the Sub-6 GHz band control channels and utilise the same mmWave band for both data and control operations.

Table 2.1 summarises the updated list of industrial standards operating at mmWave frequencies and their corresponding spectrum bands. The common characteristic of all industrial standards adopting mmWave bands is a large amount of bandwidth available that allows wireless links to achieve data rates of gigabits per second [4, 10]. Thanks to the large channel bandwidth, mmWave communication can support Enhanced Mobile BroadBand (eMBB) and Ultra-Reliable Low-Latency Communication (URLLC) services critical for 5G wireless Radio Access Network (RAN) [51, 52].

The mmWave frequencies enable compact arrays of antennas to fit into a limited space. Both mmWave-BS and UE can employ a multi-antenna array to form narrow beams that improve the beamforming gain and help overcome high frequency-dependent pathloss [53]. The combination of TX and RX beams, accurately aligned, forms a beam pair that sets a highly directive mmWave link, combining the TX and RX beamforming gains to increase the end-to-end power budget. However, this alignment is difficult to maintain and introduces several beam management challenges [54].

Unlike Sub-6 GHz networks, mmWave networks are vulnerable to the blockages on the signal path due to higher penetration losses and lower diffraction characteristics present



(a) Blockage loss for carrier frequency $f_c = 2$ GHz. (b) Blockage loss for carrier frequency $f_c = 60$ GHz.

Figure 2.1: Illustration of blockage loss obtained with the Blockage Model B indicated in 3GPP technical report 38.901 and adopted in Sec. 4.2.3 varying the blocker width (assuming infinite blocker height) and distances between blocker and Receiver (RX) (assuming the blocker in the middle between Transmitter (TX) and RX) for the carrier frequency (a) $f_c = 2$ GHz and (b) $f_c = 60$ GHz.

at the mmWave frequencies [55]. When the mmWave beam pair becomes obstructed, the channel's propagation conditions change from LoS to Non Line-of-Sight (NLoS), and the received signal power experience a significant attenuation, up to 40 dB at the 60 GHz carrier frequency [56].

In Fig. 2.1, we show an example of the blockage loss values plot obtained with the Blockage Model B indicated in [57] and adopted in Sec. 4.2.3. We place the blocker in the middle between TX and RX link, and we change the distance between TX and RX and the blocker size. To note the difference between the blockage loss values for the carrier frequency $f_c = 2$ GHz represented on the left side of Fig. 2.1, and the loss values for the carrier frequency $f_c = 60$ GHz depicted on the right side of Fig. 2.1. In the latter scenario, the loss values achieve up to 40 dB of attenuation for the cases of small TX-RX distances and blocker size larger than 4 m. We want to highlight that the blockage effect occurs in different types of scenarios that we summarise as follows:

- Scenario with mobile UE (RX) moving in a blocked region, e.g. around the corner of a building [58].
- Stationary UE scenario (RX) affected by dynamic blockages, e.g., humans, robots, or vehicles, moving around and causing temporary link blockage [56, 59].

- Vehicular scenarios in which the UE (RX) is positioned on the vehicle and the signal may be obstructed by stationary blockages, e.g., buildings and road infrastructure, and dynamic blockages, e.g. other vehicles [60].

On the other hand, while offering an abundance of frequency bands for the transmission, mmWave communication introduces several new challenges regarding practical implementation. The main limiting factors are the integration of the components in the RF chain and their power consumption [61]. RF components, e.g., Power Amplifier (PA), Low Noise Amplifier (LNA), phase shifters, Analog-to-Digital converter (ADC)/Digital-to-Analog converter (DAC) and Voltage Controlled Oscillator (VCO), require to be integrated behind each physical antenna, which is challenging to realise due to space constraints. Besides, the power consumption of these components is excessive at mmWave frequencies. For example, a commercial ADC with sampling rates in the order of 1-2 gigabits per second and high-resolution with 8-12 bits consumes around 500 mW. Therefore, a fully digital architecture like the one used for the mMIMO Sub-6 GHz system requires one RF chain for each antenna is not practical, even for small array dimensions, as the power consumption becomes prohibitively high [62].

Therefore, using an analogue-only architecture with a single RF chain is more suited for the mmWave system's implementation [47]. A phased antenna array changes the antenna phase shifters' values and steers the antenna arrays accordingly to the beam's direction. The beam is selected from a codebook containing multiple beamforming vectors that enable to cover multiple steering directions [63]. While reducing power consumption, this type of design limits the flexibility of steering the beam in only one direction at a time. As we will see in Chapter 4, mmWave systems with analog architecture adopt beam sweeping to change the beam direction and serve one UE at a time. Furthermore, the adoption of a Time-division multiple access (TDMA) scheme that divides the frame into multiple Transmission Time Interval (TTI) is well suited with analog beamforming architecture, as it allows to serve one UE in each TTI, occupying all the bandwidth.

2.1.4 Self-Backhauling

The 3GPP included in the 5G NR Release 15 a new study item, which focuses on IAB network architectures, referred to as s-BH networks in the literature. In s-BH, a SC connects through wireless backhauling to another BS that functions as an ingress node to the core network. This facilitates a large scale and cost-effective deployment of SCs. Unlike out-of-band backhauling, where the backhaul resources are assigned to a band separated from the access resources, the in-band approach requires access and backhaul to share the same spectrum resources. The backhaul can also reuse the same access hardware to provide a full integration between the two services.

In [8], 3GPP provides a list of use cases, deployment scenarios and network architecture requirements for the NR backhauling functionalities coupled with the RAN technology. The SCs are deployed to enhance the network capacity when UEs density is high and non-uniform. Additionally, SCs can provide network coverage in areas where shadow fading is severe due to the buildings' obstructing the signal. The envisioned scenarios for s-BH comprise both Sub-6 GHz and above-6 GHz spectrum for in-band backhaul operation. Several works focused on mmWave s-BH networks [64, 65], which provide a cost-effective option to extend the mmWave network coverage as the wide mmWave channel bands can simultaneously accommodate multiple backhaul and access links. At the same time, various research efforts have considered Sub-6 GHz s-BH networks in a HetNet environment [66–68], which is more suitable for providing wide-area coverage through conventional macro-cells whereas, at the same time, use wireless backhauled SCs to boost the network capacity. However, due to the scarcity of spectrum in the bands below 6 GHz, the bandwidth splitting required to serve multiple backhaul links and the inter-tier interference (from the backhaul to access links) caused by co-channel access and backhaul operations may turn out to be a significant impediment to the potential adoption of s-BH in Sub-6 GHz HetNets [68].

In [66, 67], the authors tackled the resource allocation and link management problems of s-BH networks in time and frequency domains. Moreover, in [68], the authors provided an analytical study about FD SCs, which avoids backhaul and access spectrum orthogonalisation due to the possibility of in-band bi-directional transmission and doubles the bandwidth available at the SCs for access. However, it is important to note that this option introduces

self-interference (from the access TX to the backhaul RX), which requires introducing a cancellation mechanism to prevent coverage degradation.

2.2 Fundamental Differentiators for State-of-the-art Networks

In the following subsections, we introduce the key factors that characterise the state-of-the-art networks discussed in this dissertation.

2.2.1 Heterogeneous Services

The 5G networks consider enhancing human communications and simultaneously offering wireless access to objects and machines that communicate with each other. Unlike the 4G network's services, 5G networks provide connectivity to Internet of Things (IoT) devices and autonomous smart entities, like vehicles, buildings, sensors and actuators. This new set of applications is luring sectors, such as Industry 4.0, automotive and health-care, whose main stakeholders can use the 5G networks to enable closed-loop control and remote interaction with physical objects in real-time [69]. Therefore, 5G introduces new use cases, e.g. massive Machine-Type Communication (mMTC) and URLLC, that target specifically machine-centric communications and require designing the network infrastructure to cope with new requirements, such as massive access by a large number of devices, low latency and high reliability. For instance, in the automobile sector, the exchange of safety messages between connected cars requires milliseconds latency or even smaller. Moreover, several use cases for Industry 4.0, such as actuators' control or connectivity with IoT gateways, entail reliability levels with five-nines (99.999%), e.g. gaming and Virtual Reality (VR)-Augmented Reality (AR), or even nine-nines in the most conservative cases, e.g. advanced industrial robotics [70].

2.2.2 New Spectrum Bands

5G is the first network with the flexibility to operate in multiple frequency bands and support both licensed and unlicensed scenarios for public and private networks. The 3GPP NR Standard considers two ranges of frequencies for the 5G network operation: Frequency

Range 1 (FR1), which includes Sub-6 GHz bands and Frequency Range 2 (FR2), which consist of the 28 GHz and 39 GHz bands. New spectrum bands offer high flexibility to operate in multiple frequency bands, like the NSA mode, where the mmWave bands improve the data rate while the Sub-6 GHz bands provide the control plane. Multiple frequency bands can provide more flexibility for radio spectrum operations. For example, Sub-6 GHz bands can be used more extensively outdoors to improve coverage and data rate. In contrast, mmWave bands can be adopted indoors, e.g. office spaces, shopping malls and factories, to enhance the connectivity data rate [71].

The private network deployment provides another compelling use case for industrial scenarios adopting new spectrum bands [9]. In private networks, the use of licensed spectrum is preferred because of the reliability and latency requirements needed by URLLC and mMTC services. The spectrum used by the private networks can be obtained from the regulator or subleased from the MNO; thus, private networks become independent from the public network and can apply ad-hoc security policies and control sensitive data [9]. Private networks will also be facilitated as the spectrum regulators introduce new approaches for governing spectrum usage [72]. For instance, the Federal Communications Commission (FCC) in the United States is now assigning the Citizens Broadband Radio Service (CBRS) licenses, referred to as "Innovation Band", for commercial and general use.

2.2.3 ML for 5G and Beyond Networks

Extensive network densification and the use of mMIMO and mmWave techniques require continuous updating of the network infrastructure. Network densification requires connecting with high-speed fibre backhaul every new BS deployed in the network. Similarly, mMIMO and mmWave require new specialised radio hardware to support radio transmission with these advanced techniques. This operating model has shown some limitations lately, as the MNOs require investing large capitals in deploying the network infrastructure. MNOs operating expenses (Opex) as a percentage of revenue increased from 11% in 2012 to 15% in 2017. Still, their revenue decreased by 13% during the same period [73]. Therefore, the technologies introduced for 5G will likely no longer be sufficient to continue to evolve the network infrastructure.

Table 2.2: List of state-of-the-art networks combining different 5G technology components and other enabling technologies.

	5G Technology Component				**Other Enabling Technologies**		
State-of-the-art Networks:	mMIMO Sub-6 GHz	UDN	mmWave Comm.	Self-BH	UAV	WiFi-Mesh	ML
Nokia F-Cell [18]	•	•			•		
Facebook Terragraph [19]		•	•			•	
LSTM Multi-Link Prediction for mmWave Indoor Systems [11]			•				•
Dual-band Network with Sub-6 GHz and mmWave transceivers [12]			•				•
Networks under study:							
mMIMO Self-Backhaul [20]	•	•		•			
Predictive mmWave Indoor Network [21]			•				•

On the other hands, ML tools can provide new opportunities for improving the network performance without deploying new network infrastructure as the integration of ML into the 5G networks happens by adding new software components to the network, making them more suitable for later upgrades. ML tools are envisioned as one of the essential tools to be applied across all the protocol layers to realise the future intelligent RANs [14]. Additionally, data-driven approaches supported by ML tools can optimise the management of the radio resources and increase the link robustness of the 5G and beyond network deployments [13]. Among the several case studies, in this dissertation, we focus on using the ML algorithms' prediction capabilities to manage the radio resources proactively. In Sec. 2.3.2, we will review the most relevant approaches for the mmWave links' blockage prediction problem.

2.3 State-of-the-art Networks and Methods

Accordingly to the list of differentiators explained in Sec. 2.2, the state-of-the-art network needs to be massively deployable, cost-efficient, flexible and may require proactive approaches to allocate resources. As a solution, we believe that combining 5G technologies and new enabling technologies such as ML and UAV fits this scope. As shown in Table 2.2, we identify several examples of works in the literature that follow a similar line of reasoning. In the following sections, we review the state-of-the-art network architectures for outdoor wireless backhaul and the beam recovery methods for mmWave networks. These examples served us as the basis of our work and motivated us to provide the remaining chapters' technical contributions.

2.3.1 Network Architectures for Outdoor Wireless Backhauling

Nokia F-Cell and Facebook Terragraph network architectures aim to reduce the costs of BSs deployment. They propose alternative solutions to enable the flexible deployment of new BSs through wireless backhauling instead of using fibre-optic cables. These network architectures, designed for outdoors deployments, employ mMIMO and mmWave technologies for the backhaul connectivity. In the case of the Nokia F-Cell, mMIMO is used to provide backhauling links to multiple SCs in parallel. In contrast, Facebook Terragraph employs mmWave technology to build a mesh network topology. These works' primary objective is to limit the infrastructure costs and provide services with quality comparable to the one offered by the cabled connections' services, e.g. fibre broadband.

Nokia F-Cell

The Nokia F-Cell's architecture integrates a mMIMO Hub with multiple single antenna SCs: the mMIMO Hub forms the fronthaul wireless link between the mMIMO Hub and the SCs, whereas the SCs individually serve their associated UEs [18].

The system design of the Nokia F-Cell presents several novel features. Firstly, the fronthaul link, traditionally based on the fibre-based Common Public Radio Interface (CPRI) is replaced with a wireless-based mMIMO technology installed in a mMIMO Hub. In this context, we use the term fronthaul to indicate the wireless backhaul link from the mMIMO Hub to the SC. Moreover, according to the cloud RAN (C-RAN) architecture, the SC is a Remote Radio Head (RRH) with functionalities limited to the RF operations. At the same time, the SC baseband functions reside at a central location, which manages multiple baseband units incorporated together in a pool.

Unlike other wireless backhaul solutions based on mmWave, mMIMO Sub-6 GHz supports NLoS channel propagation conditions, making the SC's deployment more flexible for locations without LoS visibility. Secondly, a novel channel training method presented in [74] and an advanced channel estimation technique shown in [75], which enable Nokia F-Cell to remove the digital signal processing operations at the SC while adopting a FDD system configuration. Thanks to this innovative design, the SC becomes equivalent to an analogue

repeater, which is extremely energy efficient to the point that a solar panel can fully power it.

Nokia F-Cell eliminates the need for wired connections (backhaul and power supply) to provide the flexibility to deploy the SC in locations not reached by the wired backhaul infrastructure or without a power source located nearby. Additionally, since the SC can be deployed in a "drop and forget" manner, the SC's transportation to the site can be facilitated by using UAVs with the capability of carrying a medium-sized weight load.

Facebook Terragraph

Facebook Terragraph project implements a mmWave mesh network to provide wireless broadband connectivity as an alternative to deliver a fibre optic connection in each building of a dense urban environment [19].

The network architecture is based on the IEEE 802.11ay standard and operates in the unlicensed 60 GHz band (V-band). Multiple Distribution Nodes (DNs) form the backhaul infrastructure, while Client Nodes (CNs) provide access connectivity to residential SCs or Wi-Fi Access Points (APs). The DNs are deployed on the street furniture, e.g. on street lamps, with a small distance. This setup helps to contrast the severe channel propagation conditions in the 60 GHz band. Indeed, one of the main challenges of using the 60 GHz band is to ensure LoS visibility between DNs.

Consequently, the DN deployment needs to be carefully planned so that each DN has multiple LoS links with other DNs if one of the links is not available. Facebook Terragraph utilises a fully automated planning procedure to identify the DN sites using Light Detection And Ranging (LiDAR) data and determines the optimal locations to deploy the DNs according to the LoS visibility conditions [76]. On the other hand, temporal dynamic factors, such as weather, foliage, and moving obstacles, can temporarily block the mmWave links. Therefore, to face this challenge, Facebook Terragraph integrates a software-based routing platform that implements a distributed control procedure to redirect the transmission packets on alternate routes in case of blockage [19].

The major innovation introduced with Facebook Terragraph is that DNs equipped with IEEE 802.11ay technology, together with Wi-Fi-mesh features, allow serving dense urban

environments with unlicensed spectrum bands. At the same time, the main limitation is that it requires multiple DNs having LoS visibility between each other to maintain the mesh network topology and offer the link's redundancy.

How to Evolve Beyond State-of-the-art Network Architectures for Outdoor Wireless Backhauling?

In a 5G network scenario, we expect the widespread utilisation of the mMIMO technology for both access and backhaul applications. Thus, in this dissertation, we initially take inspiration from the Nokia F-Cell architecture and we apply mMIMO to add more flexibility for deploying SCs and provide backhaul service in a large area reusing the same band for multiple SCs in parallel. Additionally, 5G networks are likely to support the full integration of backhauling and access services in the same spectrum bands throughout the use of the s-BH technology described in Sec. 2.1.4. A further aspect not considered by Nokia F-Cell and other works on mMIMO backhaul architectures proposed in [36, 37, 77] is the optimal use of the spectrum resources by backhaul and access legs. Therefore, it becomes interesting to study novel networks architectures that combine mMIMO and s-BH to allow flexible ultra-dense deployment of SCs. Currently, there is a lack of works in the state-of-the-art that focus on the joint access backhaul design of mMIMO-based s-BH network architecture and the strategies to follow while deploying the SCs. We will address these specific problems in Chapter 3.

2.3.2 Millimetre Wave Networks with Dynamic Blockages

In mmWave communication, the frequent transitions of the channel between LoS and NLoS states due to the blockage significantly affect the channel quality. The mmWave links become highly intermittent, degrading throughput and latency performance during the blockage events [78]. Several studies have analysed mmWave networks with dynamic blockers moving in the environment. For instance, the authors in [59] study the dynamics and the temporal correlation of the blockage in the urban environment. In [79], the authors provide a stochastic geometry blockage model that combines static and dynamic blockage effect. Then in [80], the authors analyse the impact of the blockage effect on a mmWave indoor

network. The blockage models adopted in [59, 79, 80] consider stochastic approaches. For example, the 3GPP Blockage Model A, considered in [80], account for the blockage effect that is independently evaluated for each path. A more detailed model is the 3GPP Blockage Model B, which supports spatial consistency and captures the blockage correlation in close UEs positions [57]. Our work, presented in Chapter 4, will study a mmWave network with dynamic blockers adopting this model.

As highlighted by these recent studies, in mmWave networks become essential to design control schemes that can improve mmWave links' robustness and stability. When the blockage event occurs, the beam recovery operation aims to re-establish the connection with a backup beam pair to avoid significant data rate loss. Different approaches can be taken for the beam recovery operation. In the next sections, we differentiate between control methods that rely on detecting the blockage and consequently switch to a backup beam, and we review novel ML-based approaches that predict the blockage.

Beam Recovery Methods Based on Detection

Conventional beam recovery methods based on a detection threshold search for a backup beam in alternative spatial directions from the blocked path, starting from the time instance in which the blockage occurs [81, 82]. Exhaustive search and iterative hierarchical search are the two most common techniques adopted for beam alignment to perform initial access and beam recovery. Note that the training results may stay valid after the initial access phase for a short time due to device displacement and rotation [81]. Therefore in dynamic environments, BS and UE may restart the beam search to align the TX-RX backup beam pair every time the detection triggers the beam recovery.

In the exhaustive search [83], the BS and UE scan sequentially using narrow beams with large beamforming gain all the combination between TX and RX codebook' directions. The advantage is the high beamforming gain when the beam pair align; however, the search may require a long time overhead when the TX and RX codebooks dimension are large because of the significant number of TX and RX beam combinations required to be tested. In the iterative hierarchical search [63, 84], the BS and UE perform in a first stage, an approximate scan by employing wide beams to exclude part of the spatial directions. Then, in the second

stage, they conduct a precise scan by employing narrow beams and identifying the best beamforming directions by testing only those TX and RX beam combinations not excluded during the first stage. Although the iterative hierarchical search may provide significantly less training overhead than the exhaustive search, recent studies showed that the iterative hierarchical search achieves a high probability of misalignment for low signal-to-noise ratio (SNR) and delivers performance lower than the exhaustive search [85]. The practical utilisation of the iterative hierarchical search in the SA mmWave networks, which present significant coverage issues, is still an open research problem. Thus, in this dissertation, we consider adopting a more practical exhaustive search. This is also in line with the beam failure recovery (BFR) procedure adopted by 3GPP NR Standard [86]. After a radio link failure is detected due to the blockage, the recovery procedure triggers the beam sweeping operation to identify a new beam pair with a neighbour BS through an exhaustive search and may take a long time to complete [54].

In all these methods, the blockage detection is performed real-time and all the methods establish the backup beam after detecting the blockage. More recently, new approaches based on ML tools try to predict when the blockage occurs and provide a prediction about the mmWave link-state (blocked or non-blocked) or the Reference Signal Received Power (RSRP) evolution in future time instants. We review these methods in the next section.

Blockage Prediction Methods

In this section, we review several works that use ML tools to make the blockage prediction of the mmWave link. The ML models are generally trained using the data acquired from the mmWave channel measurements or Sub-6 GHz channel measurements. Alternatively, other works use data acquired from a camera or radar, which are mounted on top of the mmWave-BS. We classify the different approaches as follows:

- *Methods using data from mmWave bands measurements:* The works [11, 87, 88] consider training ML models based on data acquired from the measurements of the mmWave-BS. The authors of [11] consider using the mmWave channel measurements as input data for a Long short-term memory (LSTM) model that predicts the received signal power of the mmWave link. In the single-cell scenario, they collect the measure-

ments from the set of LoS and NLoS paths formed with the serving BS, whereas, in the multi-cell scenario, they collect measurements from the set of LoS paths formed with all the neighbour BSs. Nevertheless, in the single-cell scenario, the simultaneous blockage of multiple mmWave paths from the serving BS is not to exclude with self-blockage, i.e. UE blocked by its own body. Moreover, in the multi-cell scenario, the RX necessitates sweeping beam towards the spatial directions of the other BSs, thus introducing a large training overhead that penalises the data transmission. The feasibility of integrating this method into network operations is still an open research problem and can be challenging because of the number of additional measurements required as input of the ML model. In [87], the authors proposed to use unsupervised online learning to predict the link-state using the Multipath Components (MPC) of the mmWave channel formed around the RX. This approach is limited to one blocker and relies on its location information, which is usually not available unless, for instance, the blocker is equipped with a localisation device. Finally, the work proposed in [88] is suited for a vehicular scenario, where the UE inside a vehicle is subject to blockage. However, this kind of scenario is outside of the scope of this dissertation.

- *Methods using data from Sub-6 GHz bands measurements:* Other spectrum bands such as those at Sub-6 GHz, allow observing the rapid variations of the mmWave channel due to the blockage tens of ms ahead of the blockage event [89]. Following this approach, the authors in [12], proposed a Deep Neural Network (DNN) model that uses the data acquired through a Sub-6 GHz channel to predict both the blockage effect and the optimal codebook's beam that maximises the achievable rate. Their results show a 90% probability of correct predictions for the mmWave link blockage state. Both these methods assume a Sub-6 GHz system that coexists with the mmWave cellular network and continuously performs the channel sensing task. This method is verified with the single-UE scenario, and it is impractical for a SA mmWave network deployment.

- *Methods using data from other sensors (camera and radar)*: Vision-aided wireless systems also referred to as vision-to-communicate (V2C) systems, use computer vision models to predict the wireless channel propagation conditions [90]. In [91], the authors

proposed a ML-based predictor of received signal power that uses visual data acquired by a camera pointing to the TX-RX path. Similarly, in [92], the authors use visual data to predict the mmWave-BS more suitable for performing the handover. Conversely, the RadMAC project is a proof-of-concept that deploys a mmWave radar on top of the mmWave AP [93]. RadMAC tracks with the radar the presence of moving objects, e.g., humans, that can eventually intersect the mmWave link, switching the beam preemptively to a secondary link during the blockage events. The applicability of the visual and radar-based systems is proved for the single-UE case, where it is possible to cover the beam path serving the UE with a wide camera angle or with the radar system. However, practical mmWave systems use beam steering with codebook-based transmission to cover multiple UEs in the sector, making it more challenging to employ visual and radar-based methods as they need to discriminate the blockage state for each beam.

How to Evolve Beyond State-of-the-Art Beam Recovery Methods?

One of the main problems of mmWave communication is the blockage sensitivity, making it challenging to operate this type of networks with mobile blockers moving around. On the other hand, we expect that 5G networks require reliable wireless connections that implicitly require stable data rate. This is especially important in indoor networks, used for connecting factories where links may require high levels of reliability. The current state-of-the-art works such as [81, 82], adopting beam recovery method based on detection, lack of results about the performance of mmWave indoor networks with dynamic blockage characteristics. In Chapter 4, we provide an analysis of this scenario and we propose a beam recovery method that considers the real-time detection of blockage to gain more insights about the stability of the UEs data rate. Moreover, the early works [11, 12, 87, 91, 92] showed promising results when applying ML tools to predict blockage. Nevertheless, these works focus on the blockage prediction problem in a single-UE scenario. The expected trend is applying the ML tools at the network level, where the challenge is the use of predictions to take actions like switching to a backup beam before the blockage event disrupts the communication with the UE. This extension requires developing ML models capable to make predictions in a

multi-UE setting. Furthermore, new methods that leverage the ML predictions to control the network procedures can be proposed and analysed when deploying the ML models online into realistic networks or a simulation environment. Our work presented in Chapter 5 tackles these challenges and proposes a beam recovery method that uses blockage predictions to perform preemptive beam switching operations in a multi-UE setting. Differently from the ML model like [11] that uses as input data non-standard SNR measurements, we use standard beam-quality measurements reported by the UEs to the BSs according to the 3GPP NR specification [94].

2.4 Summary

A major aspect of 5G networks is the integration of multiple technologies and services to provide extremely high data rate connections with contained investments for the network infrastructure while supporting low latency and high-reliable communications for emerging use cases like the ones described for the industrial scenario. At the beginning of this chapter, we presented various technologies forming the 5G networks. For each technology, we provided a high-level description of their main characteristics. We described the key differentiators that distinguish 5G and Beyond 5G networks. Then, we presented two examples of state-of-the-art networks for outdoor wireless backhauling that enable UDN at reduced infrastructure costs. Secondly, we examined beam recovery methods to be applied in mmWave communication to mitigate the blockage effect and improve the data rate stability. We explained the beam recovery methods based on a detection threshold and predictions-based approaches, which employ ML tools to predict the blockage.

In what follows, we focus on studying a novel mMIMO-based s-BH network architecture to enable UDN deployments containing the infrastructure's costs. Then, we analyse the mmWave communication's performance in an indoor scenario, considering the effect of dynamic blockers and adopting the beam recovery method based on the real-time detection of blockages. Finally, we propose a novel method based on ML tools to be applied to the beam recovery procedure of mmWave networks in a realistic multi-UE scenario.

3 Massive MIMO Self-Backhauling for Ultra-Dense Small Cell Deployments

In this chapter, we analyse the achievable User Equipment (UE) data rates of two important network architectures for 5-th Generation (5G): massive Multiple-Input-Multiple-Output (mMIMO)-based self-Backhauling (s-BH) and mMIMO for Direct Access (DA) transmission. The technical work presented in this chapter is based on our two papers [20, 95].

3.1 Introduction

The 5G wireless communication systems require to support a 1000× increase in capacity compared to the 4-th Generation (4G) networks [96]. To meet this gargantuan target, Mobile Network Operators (MNOs) can leverage different technologies, such as mMIMO and network densification, using Small Cells (SCs) Base Stations (BSs) [2, 3]. A further step towards the practical ultra-dense SC deployment is represented by s-BH, which could drive the tight integration between mMIMO and ultra-dense SC deployment, luring MNOs with the potential of achieving the desired capacity boost at a contained investment [34]. Indeed, exploiting the large number of spatial Degrees-of-Freedom (DoF) available with mMIMO to provide Sub-6 GHz in-band wireless backhauling to SCs offers multiple advantages to MNOs: *i)* avoiding deployment of an expensive wired backhaul infrastructure, *ii)* availing of more flexibility in the deployment of SCs, and *iii)* not having to purchase additional licensed spectrum, as in the case of out-of-band wireless backhauling [97].

3.1.1 Scope and Contributions

In this chapter, we analyse the end-to-end UE performance of mMIMO s-BH architecture adopting theoretical analysis and Third Generation Partnership Project (3GPP)-based system-level simulations. We consider a realistic multi-cell setup [98], where mMIMO base stations (mMIMO-BSs) provide Sub-6 GHz backhauling to a plurality of Half-Duplex (HD) SCs overlaying the macro BS coverage area. If applied to these HD systems, a s-BH network entails sharing time-and-frequency resources between radio access and backhaul links.

Two different strategies of self-backhauled SC deployments are considered, as illustrated in Fig. 3.1. We analyse a *random* deployment – where SCs are uniformly distributed over a geographical area –, and an *ad-hoc* deployment – where SCs are purposely positioned close to UEs to achieve Line-of-Sight (LoS) access links. We would like to stress that the particular choice of the ad-hoc SC deployment serves two purposes: *a)* to study the performance gains that the MNOs can obtain through a tailored SCs deployment, and *b)* to provide an upper bound of the system access performance obtained when the SC locations are coupled to the UE locations.

Additionally, we compare the above mMIMO-based s-BH network architecture to a mMIMO DA architecture accounting for the effects of different pilot reuse schemes, together with their associated overhead and reduction of the pilot contamination effect. Unlike previous works [36, 37, 77, 99], our work accounts for a pathloss model with LoS and Non Line-of-Sight (NLoS) transitions in both backhaul and access links. The closest similar work is [36], where the authors also compare mMIMO s-BH and mMIMO DA architectures' performance, but they consider channel models based only on pathloss and do not account for the pilot contamination effect on the SINR when mMIMO is used. On the contrary, we show that both the pathloss models incorporating LoS and NLoS transmissions and the pilot contamination effect impact the inter-cell interference modelling and the system performance, making our results more accurate and realistic than those presented in [36].

We finally complement this study by providing an analytical model for evaluating the average data rate of the backhaul and access links. We use 3GPP-based system-level simulations to analyse the network deployments up to one SC for each UE, whereas we use

the analytical framework when the SC density is much larger than the UE density. In this regime, numerical simulations become impractical due to the time-consuming process and the high computational complexity.

We summarise the main contributions of the chapter as follows:

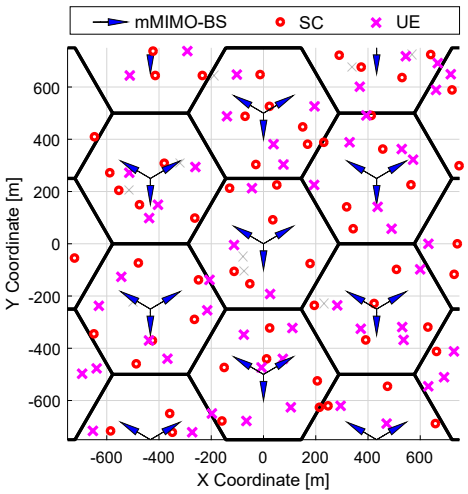
1) We propose a mMIMO-based s-BH architecture with an ad-hoc SCs deployment and we investigate the UE data rate performance changing the distance from the SC to the UE and the SC antenna pattern. Then, we analyse the same architecture with a random SCs deployment, varying the number of SCs in the network.
2) We study the effect of the backhaul/access partition variation in a mMIMO-based s-BH architecture and investigate the maximum achievable end-to-end UE data rates while optimising the backhaul/access resource allocation.
3) We compare the mMIMO s-BH and DA network architectures performance, accounting for the pilot contamination effect. We also analyse the UEs performance at different throughput percentiles to study the best network architectures for cell-edge UEs.

The remainder of this chapter is organised as follows. Section 3.2 introduces the system model on which the analysis is based; Section 3.3 presents the Downlink (DL) SINR and rate expressions of the backhaul and access links; Section 3.4 provides the analytical signal-to-interference ratio (SIR) and average rate expressions of backhaul and access; Section 3.5 presents the numerical results, and 3.6 summarises the key findings.

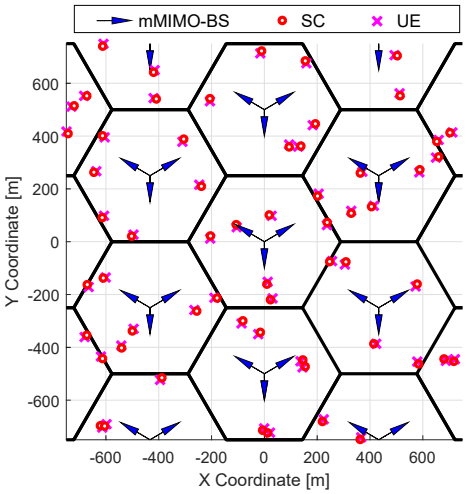
The following notation is used throughout the chapter. Capital and lower-case bold letters denote matrices and vectors, respectively, while $[\cdot]^*$, $[\cdot]^{\mathrm{T}}$ and $[\cdot]^{\mathrm{H}}$ denote conjugate, transpose, and conjugate transpose, respectively. $|\cdot|$ indicates the absolute value. The notation $\mathbb{E}[\cdot]$ indicates the expectation of a random variable (RV). We use $\mathbb{C}$ and $\mathbb{R}$ to denote complex- and real-valued numbers, respectively.

3.2 System Model

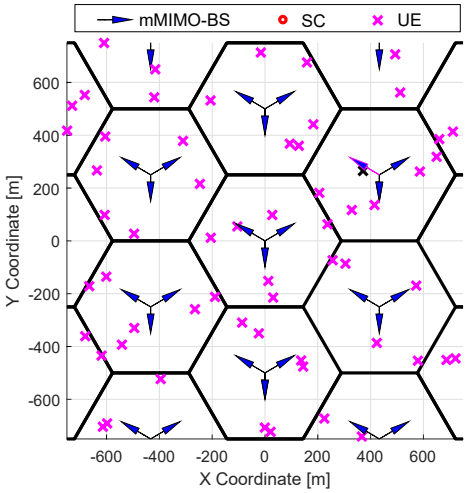
We focus on the study of the DL performance for a hexagonal grid of mMIMO-BSs equipped with a large number of antennas M and providing wireless backhaul links to (a) randomly de-



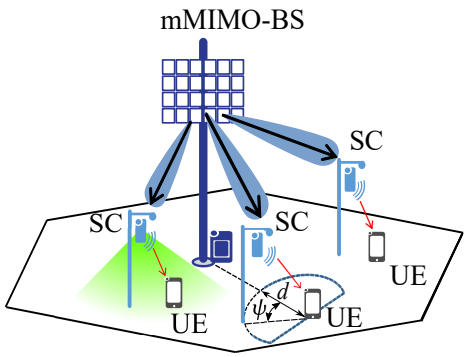
(a) s-BH architecture with random deployment of SCs



(b) s-BH architecture with ad-hoc deployment of SCs



(c) DA architecture without SCs



(d) SCs deployment in one cell of the s-BH ad-hoc architecture

Figure 3.1: (a) Top-down view of the s-BH architecture with random SCs deployment, (b) top-down view of the s-BH architecture with ad-hoc SCs deployment, (c) top-down view of the DA architecture and (d) side view of one cell of the s-BH architecture with ad-hoc SCs.

ployed self-backhauled SCs, (b) ad-hoc deployed self-backhauled SCs, or (c) directly serving UEs, as illustrated in Figs. 3.1a, 3.1b, and 3.1c, respectively. Throughout the chapter, we assume to use Sub-6 GHz frequencies, which can provide backhaul connections with a single hop to all the SCs deployed in the macro cell's coverage area [34].

For the mMIMO-BSs, we consider the hexagonal model since it represents a planned deployment scenario where the BS locations are optimised to improve the SINR performance. This deployment strategy follows the 3GPP methodology described in [98], which we take

as a reference throughout the chapter.

3.2.1 Macro Cell, Small Cell and User Topology

We denote by i, l, and k the mMIMO-BS in the i-th sector, the SC, and the UE, respectively. $\mathcal{I}$ represents the set of mMIMO-BSs deployed in the network. $\mathcal{L}_i$ and $\mathcal{L}_{i'}$ represent the set of SCs deployed per sector and connected to the i-th and i'-th mMIMO-BS, respectively, which provides the largest Reference Signal Received Power (RSRP).[1] L_i and $L_{i'}$ denote the number of SCs in the sets $\mathcal{L}_i$ and $\mathcal{L}_{i'}$, respectively. Furthermore, we denote by K_i the number of UEs in the set $\mathcal{K}_i$ randomly and uniformly distributed over the area covered by each sector. We assume that each single-antenna UE is connected with the SC (in the s-BH network) or with the mMIMO-BS (in the DA network) that provides the largest RSRP [98]. Therefore each SC serves K_l UEs in the s-BH network. Three different network deployments are presented in the following:

a) **s-BH architecture with random deployment:** Self-backhauled SCs are randomly and uniformly distributed over the mMIMO-BS cell area, as shown in Fig. 3.1a. This scenario is used as a baseline and follows the parameters specified in [98] to evaluate the relay scenario. More precisely, we consider the UE and SC antenna heights fixed at 1.5 and 5 meters above the ground, respectively, and channel models for the 3GPP Case 1 Relay scenario.

b) **s-BH architecture with ad-hoc deployment:** Self-backhauled SCs are positioned targeting nearby UE locations.[2] As shown in Fig. 3.1b and Fig. 3.1d, we model this scenario by considering SCs deployed within a distance d from the UEs and with a rotation angle $\psi \sim \mathcal{U}[-\pi/2, \pi/2]$ from the link between the UE and the closest mMIMO-BS. It is worth noting that even when the 2-D distance $d = 0$, UEs and SCs are still separated in space because the antennas are positioned at different heights, as specified in (a). Also, to limit the effect of the inter-cell interference, we replace the Patch antenna at the SC with a more directive Yagi antenna, pointing downwards

[1] We remark that a given SC deployed in the i-th sector might be connected to another mMIMO-BS i' if that provides a higher RSRP level than the mMIMO-BS i.

[2] We assume the possibility to realise this specific network deployment, for example by means of drone-BSs, where the drone-BSs can re-position themselves following the locations of UEs as suggested in [100]. Although mentioned, the drone-BSs use-case is not the focus of this chapter, and it is left for future investigation.

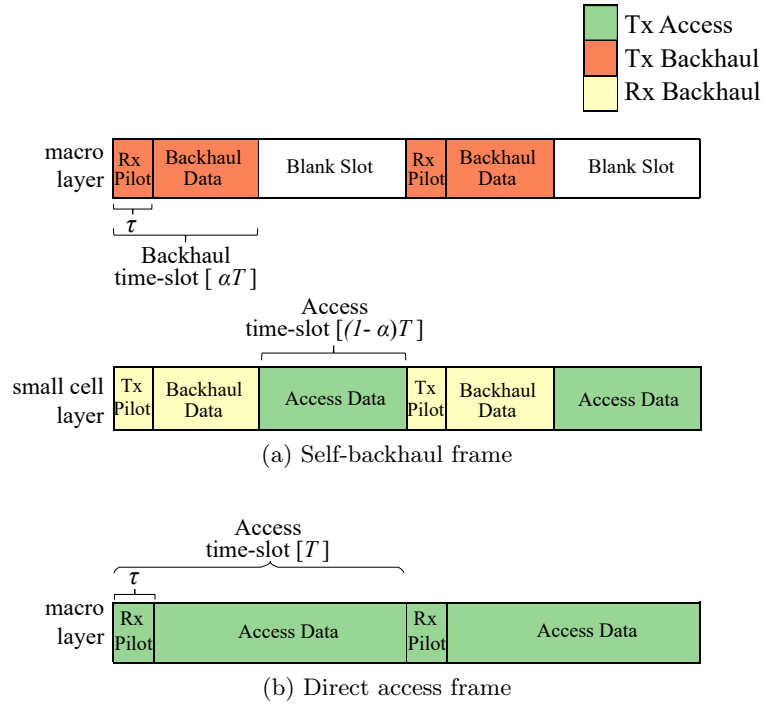


Figure 3.2: DL frame structure for (a) mMIMO s-BH with $\alpha = 0.5$, and for (b) mMIMO DA.

to the ground (as shown by the green radiation cone in Fig. 3.1d) and therefore only illuminating the closest UEs.

c) **DA architecture:** There are no self-backhauled SCs deployed, and the mMIMO-BSs serves the UEs directly, as illustrated in Fig. 3.1c.

3.2.2 Frame Structure

We consider a Time Division Duplex (TDD) system, where the time-slot duration T is used as a single scheduling unit in the time domain. As shown in Fig. 3.2a, we partition the access and backhaul DL resources through the parameter $\alpha \in [0, 1]$. The fraction α of the time-slot duration T is allocated to the backhaul links, and the fraction $1-\alpha$ of the time-slot duration T is allocated to the access links. Without loss of generality, the proposed solution described and analysed for the DL direction, can also be analysed in the Uplink (UL) direction, where there is a different sequence of access and backhaul operations. In the frequency domain, we divide the system bandwidth BW into Q_t Resource Blocks (RBs), and we allocate all the RBs to the backhaul links or the access links. We make the following assumptions in considering the partition of backhaul and access time-slots among the SCs and UEs:

- During the backhaul time-slots, all the associated SCs are served by the mMIMO-BS i with M antennas, and we use the same value of α for all the SCs. The mMIMO-BSs precode the backhaul signals towards the single-antenna SCs, which are spatially multiplexed in the same time-frequency resources. We consider to have M DoF for a mMIMO-BS with M antennas [101], and we assume $M > \mathbb{E}[L_i]$ by dimensioning the number of antennas M during the mMIMO-BS deployment planning. In the case that L_i may exceed the number of antennas M, a scheduler ensures to serve only M SCs randomly picked in the set $\mathcal{L}_i$. We also observed in our experiments that $\Pr[L_i > M]$, i.e. the probability that the RV L_i exceeds the number of DoF M, decays rapidly for large M, and is equal to zero for the set of system parameters adopted in our simulations.
- During the access time-slots, the SCs schedule their connected UEs by using a Round Robin (RR) mechanism, which equally distributes the available Q_t RBs among its UEs.

Fig. 3.2b shows the frame structure used for the DA setup, where the entire time-slot duration T is allocated to the access links. In each time-slot, the mMIMO-BS precode the access signals, and the UEs are spatially multiplexed reusing the entire system bandwidth BW. We assume $M > \mathbb{E}[K_i]$ due to the system design. When $K_i > M$, a scheduler serves only M UEs randomly picked in the set $\mathcal{K}_i$. The same observation made above for the backhaul applies with K_i in place of L_i.

Figs. 3.2a and 3.2b also show the fraction τ of the time-slots dedicated to transmitting the UL pilot sequences, used for the Channel State Information (CSI) acquisition. Details about the CSI acquisition procedure will be discussed in Sec. 3.2.4.

3.2.3 Channel Model

We define as $\mathbf{h}_{il} = [h_{il1}, \ldots, h_{ilM}]^{\mathrm{T}} \in \mathbb{C}^M$ the propagation channel between the l-th single-antenna receiver (SC in the mMIMO s-BH network and UE in the mMIMO DA) and the M antennas of the i-th mMIMO-BS. The composite channel matrix between the i-th mMIMO-BS and the receivers in the i'-th cell is represented by $\mathbf{H}_{i,i'} = [\mathbf{h}_{i1} \cdots \mathbf{h}_{iL_{i'}}] \in \mathbb{C}^{M \times L_{i'}}$, where we omit the subscript q indicating the q-th RB of the channel matrix for

notational convenience. Moreover, for the mMIMO s-BH architecture, we define the single-input single-output (SISO) channel between the l-th SC and the k-th UE in the q-th RB as $g_{lkq} \in \mathbb{C}$.

The channel coefficients $h_{ilm} = \sqrt{\beta_{il}}\widetilde{h}_{ilm}$ and $g_{lkq} = \sqrt{\beta_{lk}}\widetilde{g}_{lkq}$ account for both the effects of the large-scale fading and the small-scale fading components:

- We model the large-scale fading components $\beta_{il}, \beta_{lk} \in \mathbb{R}^+$ according to the 3GPP Case 1 Relay scenario [98]. For a given link, the model decides whether the channel propagation conditions are LoS or NLoS, considering a distance-dependent LoS probability function, and uses log-normal distributed shadowing with different values of standard deviation. Because of its slow-varying characteristic, the large-scale fading does not change rapidly with time, and it can be assumed constant over the observation time-scale of the network.
- We model the small-scale fading components $\widetilde{h}_{ilm}, \widetilde{g}_{lkq} \in \mathbb{C}$, which result from multi-path as Rician fast-fading, according to the 3GPP spatial channel model for MIMO simulations [102], assuming a K-factor dependent on the distance between transmitter and receiver.

Throughout the chapter, we assume a block-fading channel model, where the channel vectors of the access paths, i.e. the channel between each antenna at the mMIMO-BS and the UE, and the channel between the SC and the UE, remain constant for a frequency-time block corresponding to one time-slot T, and one RB [28, 103, 104]. In line with the Long Term Evolution (LTE) numerology [105], we consider the time-slot duration is 1 ms, and the bandwidth of the RB is 180 kHz. Differently, we consider that the channel vectors of the backhaul path, i.e. the channel between each antenna at the mMIMO-BS and the SC, is constant for a period $T_{BH} \gg T$ due to the static position of the SCs. Moreover, we adopt a horizontal Uniform Linear Array (ULA) for the mMIMO-BS with equally spaced antennas and Jakes correlation model between antenna pairs [106].

3.2.4 Massive MIMO CSI Acquisition

To calculate the DL precoder of the mMIMO-BS, we consider that the channel is estimated through UL pilot sequences, assuming UL/DL channel reciprocity [3]. We also consider

that the SCs or UEs associated with the same mMIMO-BS have orthogonal pilot sequences, and define the pilot codebook with the matrix $\mathbf{\Phi}_i = [\boldsymbol{\phi}_{i1} \cdots \boldsymbol{\phi}_{iL_i}]^{\mathrm{T}} \in \mathbb{C}^{L_i \times B}$, which satisfies $\mathbf{\Phi}_i \mathbf{\Phi}_i^{\mathrm{H}} = \mathbf{I}_{L_i}$. Here, the l-th sequence is given by $\boldsymbol{\phi}_{il} = [\phi_{il1}, \ldots, \phi_{ilB}]^{\mathrm{T}} \in \mathbb{C}^B$, and B denotes the pilot codebook length. Note that $L_i \leq B$, i.e. the number of SCs trained by the mMIMO-BSs in the backhaul time-slots is limited by the maximum number of orthogonal pilot sequences. Similarly, $K_i \leq B$, i.e. the number of UEs trained by the mMIMO-BSs in the access time-slots is limited by the maximum number of orthogonal pilot sequences. In case $L_i > B$ or $K_i > B$, the scheduler assigns the pilot sequences to only B SCs or UEs, randomly selected in the sets $\mathcal{L}_i$ and $\mathcal{K}_i$.[3]

The matrix $\mathbf{Y}_i \in \mathbb{C}^{M \times B}$ of pilot sequences received at the i-th mMIMO-BS can be expressed as [107]

$$\mathbf{Y}_i = \sqrt{P_{il}^{\mathrm{ul}}} \sum_{i' \in \mathcal{I}} \mathbf{H}_{i,i'} \mathbf{\Phi}_{i'} + \mathbf{N}_i, \tag{3.1}$$

where P_{il}^{ul} is the power used for UL pilot transmission by the l-th device, located in the i-th sector, and $\mathbf{N}_i \in \mathbb{C}^{M \times B}$ represents the additive noise matrix, whose entries are modelled as independent and identically distributed complex Gaussian RVs with variance σ^2.

Let us denote by $\mathbf{H}_i = [\mathbf{h}_{i1} \cdots \mathbf{h}_{iL_i}] \in \mathbb{C}^{M \times L_i}$ the channel between the i-th mMIMO-BS and the associated devices. During the UL training phase, the mMIMO-BS obtains an estimate of $\mathbf{H}_i$ by correlating the received signal with a known pilot matrix $\mathbf{\Phi}_i$. Let us define $\mathcal{P} \subseteq \mathcal{I}$ as the subset of sectors, whose devices share identical pilot sequences with the devices served by the i-th mMIMO-BS. In line with other studies [107, 108], we adopt the least-squares (LS) channel estimation since we consider a practical system, where the mMIMO-BS does not have perfect knowledge of the cross-cell channel statistics. The resulting LS channel estimation can be expressed as [107]

$$\widehat{\mathbf{H}}_i = \frac{1}{\sqrt{P_{il}^{\mathrm{ul}}}} \mathbf{Y}_i \mathbf{\Phi}_i^{\mathrm{H}} = \mathbf{H}_i + \sum_{i' \in \mathcal{P}} \mathbf{H}_{i,i'} + \frac{1}{\sqrt{P_{il}^{\mathrm{ul}}}} \mathbf{N}_i \mathbf{\Phi}_i^{\mathrm{H}}. \tag{3.2}$$

The first, second and third terms on the right-hand side of Eq. (3.2) represent the estimated channel, a residual pilot contamination component and the noise after the pilot sequence

[3]The number of pilot sequences is in most of the cases sufficient to cover all the UEs or SCs in a cell since we focus on the Enhanced Mobile BroadBand (eMBB) service and we address the conventional use case of human-type communications, where there are tens of UEs per cell and up to one active SC for each UE.

Table 3.1: Pilot allocation schemes in mMIMO s-BH and mMIMO DA architectures.

Architecture	**Backhaul links**	**Access links**
mMIMO s-BH	pilot reuse for all the sectors (No pilot contamination)	No pilot allocation
mMIMO DA	-	pilot reuse 1 (R1) pilot reuse 3 (R3)

correlation, respectively. The use of the same set of orthogonal pilot sequences among different sectors leads to the well-known *pilot contamination* problem, which can severely degrade the performance of mMIMO systems [3, 45].

Like [45], we consider that each uplink training symbol τ contains up to 16 orthogonal pilot sequences, and we consider its duration equal to one Orthogonal Frequency Division Multiplexing (OFDM) symbol period of the LTE standard [105]. The pilot sequences are assigned to SCs and UEs, in the case of mMIMO for backhaul and mMIMO for access, respectively.

In the mMIMO s-BH architecture, due to the static position of the SCs, the backhaul channel vectors remain constant for a longer time than the access channel vectors, and there is no need to update the CSI every time-slot. This allows multiplexing in time, over separate backhaul time-slots, the pilots transmitted by the SCs in different sectors. Therefore, we assume that no pilot contamination is present in the system, and we account for a pilot overhead τ over T, equivalent in our system model to 7% of the time-slot duration.

In the mMIMO DA architecture, we use two pilot allocation schemes [45]:

- In *pilot reuse* 1 *(R1)* the same set of pilot sequences is reused in all the network sectors, thus introducing strong pilot contamination. The pilot overhead to account for is τ over T, equivalent in our system model to 7% of the time-slot duration.
- In *pilot reuse* 3 *(R3)* we consider coordination between the same cell sites, avoiding using the same pilot sequence from the adjacent sectors. This reduces the effect of the pilot contamination at the expense of increasing the pilot overhead, which is $3 \times \tau$ over T, equivalent in our system model to 21% of the time-slot duration.

In Table 3.1 we summarise the different pilot allocation schemes used in mMIMO s-BH and mMIMO DA architectures.

3.3 Downlink SINR and User Rate

In this section, we present the formulation for the two-hop DL data rate in the s-BH network, which comprises the formulation for the mMIMO backhaul and the SC access SINRs and data rates. Moreover, we include the conventional formulation for the data rates in mMIMO DA network.

3.3.1 Massive MIMO Backhaul Transmission

The i-th mMIMO-BS uses the precoding matrix $\mathbf{W}_i = [\mathbf{w}_{i1} \cdots \mathbf{w}_{iL_i}] \in \mathbb{C}^{M \times L_i}$ to serve its connected SCs during the backhaul time-slot. In this chapter, we consider that $\mathbf{W}_i$ is computed based on the zero-forcing (ZF) criterion as [109][4]

$$\mathbf{W}_i = \widehat{\mathbf{H}}_i \left(\widehat{\mathbf{H}}_i^{\mathrm{H}} \widehat{\mathbf{H}}_i \right)^{-1} \mathbf{D}_i{}^{\frac{1}{2}}. \tag{3.3}$$

Here, the diagonal matrix $\mathbf{D}_i = \mathrm{diag}\left(\rho_{i1}, \rho_{i2}, \ldots, \rho_{iL_i}\right)$ is chosen to distribute the total DL power P_i^{dl} equally among the L_i receivers. In Eq. (3.3), ρ_{il} represents the power allocated to the l-th receiver located in the i-th sector, and $\mathrm{Tr}\{\mathbf{D}_i\} = P_i^{\mathrm{dl}}$, where $\mathrm{Tr}\{\mathbf{D}_i\}$ is the trace of the matrix $\mathbf{D}_i$.

Under the assumption that each SC has perfect CSI available, the DL SINR of the l-th stream transmitted by the i-th mMIMO-BS can be expressed as

$$\mathrm{SINR}_{il}^{\mathrm{B}} = \frac{\rho_{il} |\mathbf{h}_{il}^{\mathrm{H}} \mathbf{w}_{il}|^2}{\sum\limits_{\substack{j \in \mathcal{L}_i \\ j \neq l}} \rho_{ij} |\mathbf{h}_{il}^{\mathrm{H}} \mathbf{w}_{ij}|^2 + \sum\limits_{\substack{i' \in \mathcal{I} \\ i' \neq i}} \sum\limits_{j \in \mathcal{L}_{i'}} \rho_{i'j} |\mathbf{h}_{i'l}^{\mathrm{H}} \mathbf{w}_{i'j}|^2 + \sigma_n^2}. \tag{3.4}$$

The numerator of Eq. (3.4) contains the power of the signal intended for the l-th receiver, and the denominator includes the intra-cell interference from the serving i-th mMIMO-BS, the inter-cell interference from other mMIMO-BSs, and the power of the thermal noise at the SC receiver σ_n^2.

The corresponding DL backhauling rate at the l-th SC receiver can therefore be expressed

[4] As shown in [22], among the linear methods for the mMIMO precoder, ZF and minimum mean square error (MMSE) perform better than Maximum Ratio Combining (MRC). However, ZF provides a better balance between performance and complexity than MMSE.

as

$$R_{il}^{\mathrm{B}} = \alpha \left(1 - \frac{\tau}{T}\right) BW \log_2 \left(1 + \mathrm{SINR}_{il}^{\mathrm{B}}\right), \tag{3.5}$$

where α, as indicated before, represents the fraction of time-slots allocated to the backhaul links.

3.3.2 Small Cell Access Transmission

We recall from the channel model that g_{lkq} denotes the SISO channel between the l-th SC and the k-th UE corresponding to the q-th RB. The DL SINR of the k-th UE served by the l-th SC in RB q can be expressed as

$$\mathrm{SINR}_{lkq}^{\mathrm{A}} = \frac{P_l^{\mathrm{dl}} |g_{lkq}|^2}{\sum\limits_{i \in \mathcal{I}} \sum\limits_{\substack{l' \in \mathcal{L}_i \\ l' \neq l}} P_{l'}^{\mathrm{dl}} |g_{l'kq}|^2 + \sigma_{n_2}^2}, \tag{3.6}$$

where P_l^{dl} and $P_{l'}^{\mathrm{dl}}$ are the transmit powers on the RB of the l-th and l'-th SCs, respectively, and $\sigma_{n_2}^2$ denotes the thermal noise power at the UE receiver.

Therefore, the corresponding DL access rate for UE k served by SC l can be expressed as

$$R_{lk}^{\mathrm{A}} = (1 - \alpha) \frac{BW}{Q_t} \sum_{q=1}^{Q_t} x_q^k \log_2 \left(1 + \mathrm{SINR}_{lkq}^{\mathrm{A}}\right), \tag{3.7}$$

where $x_q^k = 1$ if the q-th RB is assigned to the k-th UE, and $x_q^k = 0$ otherwise.

The potential aggregated DL access rate provided by the l-th SC is $R_l^{\mathrm{A}} = \sum_{k=1}^{K_l} R_{lk}^{\mathrm{A}}$. However, the actual aggregated DL access rate provided by the l-th SC cannot be larger than the backhaul DL rate, which entails that $R_l^{\mathrm{A}} \leq R_{il}^{\mathrm{B}}$, $\forall l \in \mathcal{L}_i$, and $\forall i \in \mathcal{I}$. In this chapter, we assume that the backhaul capacity is equally divided between the K_l UEs served by the l-th SC.[5] Therefore, the resulting end-to-end access rate for the k-th UE can be expressed as

$$R_{ilk} = \min \left(\frac{R_{il}^{\mathrm{B}}}{K_l}, R_{lk}^{\mathrm{A}} \right), \tag{3.8}$$

[5]The assumption of equally distributed backhaul capacity among the UEs might become a drawback for the end-to-end rates, when UEs served by the same SC have different rate requirements in the access links, and in this case, the partition of the backhaul resources among the UEs could be designed according to their demands. This access-based partition of the backhaul resources among the UEs is not the focus of this chapter, and its study in the context of s-BH architecture is left for future work.

where the minimum is computed between the backhaul rate obtained from Eq. (3.5), divided by the number of UEs served by the l-th SC, and the access rate obtained from Eq. (3.7).

3.3.3 Massive MIMO Direct Access Transmission

In contrast to s-BH setups, mMIMO systems providing DA dedicate all their time resources to multiplex DL data streams to the UEs. Thus, the DL access rate of the k-th UE served by the i-th mMIMO-BS can be expressed as

$$R_{ik}^{\mathrm{DA}} = \left(1 - \frac{\tau}{T}\right) BW \log_2\left(1 + \mathrm{SINR}_{ik}^{\mathrm{DA}}\right), \tag{3.9}$$

where the estimated channel matrix between the i-th mMIMO-BS and its connected UEs can be expressed as $\widehat{\mathbf{H}}_i^{\mathrm{DA}} = [\widehat{\mathbf{h}}_{i1} \cdots \widehat{\mathbf{h}}_{iK_i}] \in \mathbb{C}^{M \times K_i}$, and is plugged into Eq. (3.3) to subsequently derive the access DL SINR in Eq. (3.4), which assumes that each UE has perfect CSI available, and the DA rate in Eq. (3.9).

3.4 Analytical SIR and Average Backhaul and Access Rates

In the following, we present a tractable formulation to model the mMIMO s-BH network, which approximates the backhaul and access SINR and data rate expressions.

We recall from Sec. 3.2 that the mMIMO-BS locations are distributed on a hexagonal grid with density $\lambda_a = 3(\frac{3\sqrt{3}}{2}R^2)^{-1}$, where $R = \frac{d_{ISD}}{\sqrt{3}}$ is the outer sector radius of the hexagonal site, and d_{ISD} is the Inter Site Distance (ISD) between two cell sites. For the random deployment, we consider that the locations of SCs and UEs are drawn from two independent SPPPs Ω_b and Ω_u with densities λ_b and λ_u, respectively. For the ad-hoc deployment, we consider that the SC locations are coupled to the UE locations drawn from a SPPPs Ω_u, and we position the SCs at a fixed distance d to the UE locations. The mean numbers of SCs and UEs in the finite area of the sector are obtained as $\mu_b = \lambda_b/\lambda_a$ and $\mu_u = \lambda_u/\lambda_a$, respectively [110]. Moreover, when considering a dense deployment of SCs, there is a high probability that there are unloaded SCs, i.e. without UEs associated with them [111]. We capture this effect in the analysis by considering an activation probability

for the SC, which can be approximated as [112]

$$p_a \approx 1 - \left(1 + \frac{\lambda_u}{3.5\lambda_b}\right)^{-3.5}, \tag{3.10}$$

where the expression $\left(1 + \frac{\lambda_u}{3.5\lambda_b}\right)^{-3.5}$ approximates the average void probability for a typical Voronoi cell [112]. Consequently, the locations of the active SCs can be approximated as SPPP $\widetilde{\Omega}_b$ with density $\widetilde{\lambda}_b = p_a\lambda_b$ derived by thinning the SC process Ω_b [112]. Differently, in the ad-hoc deployment, we assume $p_a = 1$, meaning that each SC is active and serves the nearby UE. The mean number of active SCs in a sector is obtained as $\widetilde{\mu}_b = \widetilde{\lambda}_b/\lambda_a$.

In the following expressions, SIR is used to approximate the SINR, since in the Sub-6 GHz bands, with a system design which assures signal coverage, the system operates in interference-limited conditions, where the power of received interference dominates the denominator of the SINR.

3.4.1 Average Rates of Massive MIMO Backhaul Transmission

We now provide an analytical model for evaluating the average data rate of the backhaul links, given the SIR of a typical SC and the spatial distribution of SCs in the sector. Inspired by [113], we treat the SCs as UEs, and we extend the model proposed in [114] to analyse the SIR by considering: *i)* the effects of antenna directionality and sectorization, captured with the horizontal and vertical antenna patterns, and by modelling the co-site interference component; *ii)* the effect of the beamforming gain due to the mMIMO precoding as proposed in [77]. We make the following assumptions in our analytical backhaul model:

- For simplicity, the backhaul channel is statistically modelled by considering only the large-scale fading component, excluding shadowing statistics.
- We assume LoS propagation channel conditions from the serving mMIMO-BS to the SC (and from the co-site interfering mMIMO-BSs to the SC), to reflect the characteristics of the backhaul link, which tends to have dominant LoS conditions between SCs and the antennas of the nearest mMIMO-BSs [98]. On the other hand, we assume all the mMIMO-BSs from the surrounding interfering sites to be in NLoS.
- We assume that the co-channel interference from the mMIMO-BS to other served SCs

can be reasonably neglected since we adopt the ZF precoder. For consistency with the assumptions made in Sec. 3.2.4 and in Sec. 3.3.1, we consider no pilot contamination in the backhaul links, and we consider that each SC has perfect CSI available.

For the analysis that follows, r and θ denote two independent RVs, which define the distance and the angle from the SC to the serving mMIMO-BS. Note that r and θ are distributed with uniform probability density functions (pdfs) $f_R(r)$ and $f_\Theta(\theta)$ in the interval $[r_{\min}, \frac{d_{ISD}}{2}]$ and $[-\pi/3, +\pi/3]$, respectively, where the distance $r_{\min}$ denotes the minimum distance between the mMIMO-BS and the SC.[6] By convention, $\theta = 0$ indicates the bore-sight direction in the first sector, and such sector is denoted as $s = 1$ for each hexagonal cell formed by S sectors.

The SIR of a typical SC associated with the mMIMO-BS is modelled as

$$\mathrm{SIR}^{\mathrm{B}}(r,\theta) \approx \frac{(M - \widetilde{\mu}_b + 1)}{\widetilde{\mu}_b} \frac{P_i^{\mathrm{dl}} G_a G_{\mathrm{V}}(r) G_{\mathrm{H},1}(\theta) \beta^{\mathrm{L}}(r)}{I_1(r,\theta) + I_2(r)}, \tag{3.11}$$

where the multiplying factor $(M - \widetilde{\mu}_b + 1)/\widetilde{\mu}_b$ expresses the beamforming gain from mMIMO precoding.[7] G_a, $G_{\mathrm{V}}(r)$ and $G_{\mathrm{H,s}}(\theta)$ are the antenna gains of the single mMIMO-BS element, the vertical (V), and the horizontal (H) antenna patterns, respectively. $\beta^{\mathrm{L}}(r) = A^{\mathrm{L}} r^{-\eta^{\mathrm{L}}}$ is the pathloss between mMIMO-BS and SC, where A^{L} and η^{L} indicate the frequency-dependent pathloss factor and the pathloss exponent for the backhaul link in LoS condition, respectively, and $I_1(r,\theta)$ and $I_2(r)$ are the co-site and inter-site interference components, respectively. The vertical antenna pattern is defined as [98]

$$G_{\mathrm{V}}(r)|_{\mathrm{dBi}} = -\min\left(12\left(\frac{\arctan\frac{\delta_a}{\sqrt{r^2-\delta_a^2}} - \zeta_{\mathrm{tilt}}}{\zeta_{HP}}\right)^2, F_v\right), \tag{3.12}$$

where δ_a is the difference in antenna heights between the mMIMO-BS and the SC, ζ_{tilt} is the mechanical downtilt, ζ_{HP} is the half-power vertical beamwidth, and F_v is the vertical

[6] We observed that for the value of $r_{\min}$ that we use in our simulations, the realisations of the (modified) spatial process could be statistically considered as if they were derived from a homogeneous SPPP since we tested the spatial randomness with the Ripley's K-function [115], and we verified that the theoretical values of the SPPP are contained within the region defined by the lower and upper envelopes given by the empirical results.

[7] Only the RSRP SCs are spatially multiplexed in the backhaul time-slots, since those inactive are not required to backhaul the UEs data.

front-back ratio. Similarly, the horizontal antenna pattern is defined as [98]

$$G_{\mathrm{H,s}}(\theta)|_{\mathrm{dBi}} = -\min\left(12\left(\frac{\theta-(s-1)2\pi/3}{\theta_{HP}}\right)^2, F_h\right), \tag{3.13}$$

where θ_{HP} is the half-power horizontal beamwidth, and F_h is the horizontal front-back ratio. In Eq. (3.11), $I_1(r,\theta)$ is represented as

$$I_1(r,\theta) = P_{i'}^{\mathrm{dl}} G_a G_{\mathrm{V}}(r) \sum_{s=2}^{S} G_{\mathrm{H,s}}(\theta)\beta^{\mathrm{L}}(r), \tag{3.14}$$

and $I_2(r)$ is approximated as [114]

$$I_2(r) \approx \frac{2\pi\lambda_a P_{i'}^{\mathrm{dl}} G_a G_{\mathrm{V}}(2R_c - r)\overline{G}_{\mathrm{H}} A^{\mathrm{NL}}}{\eta^{\mathrm{NL}} - 2} \times \left((2R_c - r)^{2-\eta^{\mathrm{NL}}} - (R_b - r)^{2-\eta^{\mathrm{NL}}}\right), \tag{3.15}$$

where $R_b = \frac{3}{2}d_{ISD}$ denotes the network boundary, $R_c = \frac{d_{ISD}}{2}$ is the inner sector radius, $\overline{G}_{\mathrm{H}} = \int_0^{2\pi} \sum_{s=1}^{S} G_{\mathrm{H,s}}(\theta)d\theta$ is the average horizontal antenna gain, and η^{NL} and A^{NL} are the pathloss exponent and the frequency dependent pathloss factor for the backhaul link in NLoS condition, respectively. Eq. 3.15 approximates the integration area used to calculate the inter-cell interference from other macro-BSs assuming a ring formed by two concentric discs centred at the origin and with radii of $2R_c$ and R_b. We refer to [114] for the validation of this model. Finally, the average SC data rate for backhaul transmission can be expressed as

$$\overline{R^{\mathrm{B}}} = \alpha\, BW \times \int_{-\frac{\pi}{3}}^{\frac{\pi}{3}} \int_{r_{\min}}^{R_c} \log_2\left(1 + \mathrm{SIR}^{\mathrm{B}}(r,\theta)\right) f_R(r) f_\Theta(\theta)\, dr\, d\theta. \tag{3.16}$$

Therefore, the average SC data rate results can be computed by numerical integration of Eq. (3.16).

3.4.2 Average Rates of Small Cell Access Transmission

Inspired by the stochastic geometry analysis presented in [116], we now provide an analytical model for evaluating the access SIR of a typical UE at the origin, and its access average DL data rate. Similarly to [30, 117], we consider that the LoS and NLoS pathloss characteristics' impact the SIR modelling. We use the same LoS probability function as in [117]; however, we include in the inter-cell interference computation the density of the active SCs, i.e. those

with UEs associated.

We make the following assumptions in our analytical access model:

- We assume that each UE connects to the nearest SC.[8] For the random deployment, since Ω_b and Ω_u are two independent SPPPs, we assume that x, i.e. the distance between UE and the serving SC is a RV Rayleigh distributed with pdf $f_X(x) = 2\pi\lambda_b x \exp(-\lambda_b \pi x^2)/\exp(-\lambda_b \pi \delta_b^2)$, where δ_b denotes the difference between the SC and UE heights [116]. Differently, for the ad-hoc deployment, we assume $x = \sqrt{d^2 + \delta_b^2}$.
- The propagation channels are represented with a combination of distance-dependent pathloss and multi-path fading, distributed as Rayleigh with an exponential power distribution $|g|^2 \sim \exp(1)$.[9]
- We adopt a probabilistic LoS channel model for the inter-cell interference, with a LoS probability expressed as [117]

$$\Pr^{\mathrm{L}}(x) = \exp(-(x/D)^2), \tag{3.17}$$

  where the parameter D is set to approximate the LoS probability of the SC-UE 3GPP model [117].

The SIR of a typical UE at the origin associated with the SC is modelled as

$$\mathrm{SIR}^{\mathrm{A}}(x) \approx \frac{P_l^{\mathrm{dl}} G_b |g|^2 \beta^{\mathrm{L}}(x)}{I_{\mathrm{agg}}}, \tag{3.18}$$

where G_b is the SC antenna gain, $|g|^2$ is the multi-path channel gain, and I_{agg} is the aggregate inter-cell interference.

We now use (3.18) to provide an expression for the rate coverage probability, which defines the probability that the UE rate is higher than a minimum target R_{th}. This probability can be expressed as $\Pr\left[\mathrm{SIR}^{\mathrm{A}}(x) > \gamma_a\right]$, where $\gamma_a = 2^{R_{th}\mu_l} - 1$ depends on R_{th}, and we approximate the number of UEs associated with the SC serving the typical UE at the origin with the corresponding mean μ_l, given by $1 + 1.28\lambda_u/\lambda_b$ [118]. We later show in Fig. 3.6b

[8] This assumption holds in the considered range of SC densities, for which the probability that the closest SC is in LoS is very high, as shown by the simulation results in Fig. , due to the proximity of SC to UE.

[9] Shadowing statistics are neglected in the analytical model, although a more comprehensive framework can incorporate this effect in the distribution of the UE distances.

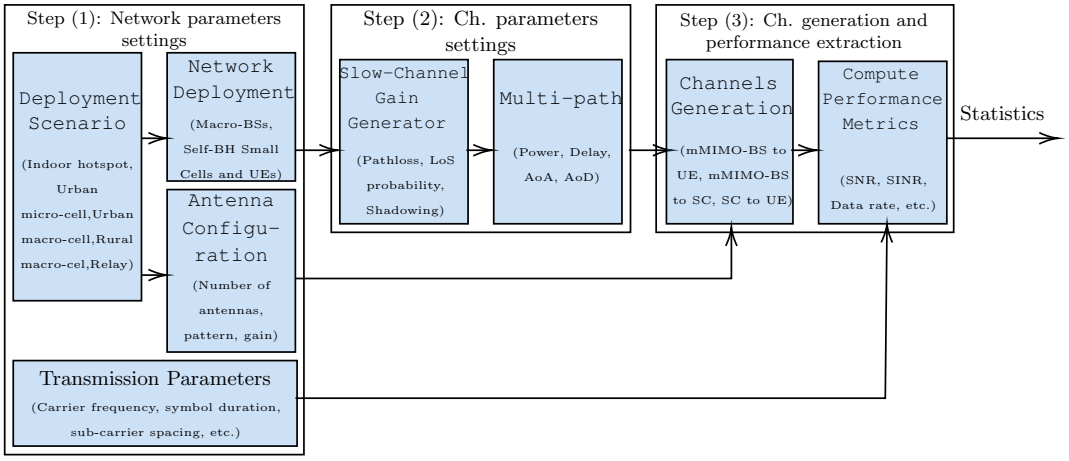


Figure 3.3: Illustration of the main components forming the system-level simulator.

that for a typical SC, the mean of the distribution of the UEs number expressed as $\lambda_u/\widetilde{\lambda}_b$ [119], matches the numerical results.

The expressions used to evaluate the rate coverage probability are included in Appendix A. Thus, the average UE data rate for access transmission can be expressed as

$$\overline{R^{\mathrm{A}}} = (1-\alpha)\, BW \times \int_0^{+\infty} \int_{\delta_b}^{+\infty} \Pr\left[\mathrm{SIR}^{\mathrm{A}}(x) > \gamma_a\right] f_X(x)\, dx\, d\gamma_a. \tag{3.19}$$

Therefore, the average UE data rate results can be computed by numerical integration of Eq. (3.19).

In the next section, we will use this model to complement the 3GPP-based system-level simulations' results.

3.5 Simulations and Numerical Results

In this section, we evaluate the performance of mMIMO s-BH and DA networks using 3GPP-based system-level simulations and mathematical analysis.

3.5.1 System-level Simulator Description

As illustrated in Fig. 3.3, the data rate statistics are computed after three major steps that are modelled following the methodology described in the 3GPP technical specification [98], and are described as follows:

(1) We set the simulation parameters following the 3GPP Case 1 Relay scenario described in [98], changing where necessary the parameters to study the two s-BH network topologies described in Sec. 3.2.1. Table 3.2 contains the full list of relevant parameters used for the network deployment, antenna configuration and transmission.

(2) We set the channel parameters adopting the LoS and NLoS pathloss exponents $\eta^{\mathrm{L}} = 2.35$ and $\eta^{\mathrm{NL}} = 3.63$ for the backhaul links, and we consider $\eta^{\mathrm{L}} = 2.09$ and $\eta^{\mathrm{NL}} = 3.75$ for the access links, as described in [98, Tab. A.2.1.1.2-3]. For each link, we use the corresponding LoS probability function proposed in [98, Tab. A.2.1.1.2-3]. To simulate the backhaul links, we account for the SC site planning correction factor, which affects the pathloss and the LoS probability as indicated in [98, Tab. A.2.1.1.4-2]. To simulate the access links, we assume cross-correlated shadowing, with correlation coefficient $\rho = 0.5$ between the UE and the SC locations [98, Tab. A.2.1.1.2-3]. In the simulations, we consider a Rician fading model, and we characterise the Rician K-factor with the model: $K[\mathrm{dB}] = 13 - 0.03\mathrm{r}$ in dB, where r is the distance between transmitter and receiver in meters [102].

(3) In this step, we generate the channel gains (composed by pathloss, shadowing and multi-path fading) for all serving and interfering radio links between each SC and the UEs, as well as between each mMIMO-BS and all SCs. We collect statistics for different network realisations, each with independent deployments of UEs and SCs. Subsequently, we measure the performance in terms of the Cumulative Distribution Functions (CDFs) of the backhaul SC rates in Eq. (3.5), of the access UE rates in Eq. (3.7), and of the end-to-end UE rates in Eq. (3.8).

To compare mMIMO s-BH against mMIMO DA architectures, we also simulate the links between mMIMO-BSs and UEs, and compute the resulting rates in Eq. (3.9). In the mMIMO DA network, we adopt the LoS probability function and the corresponding exponents $\eta^{\mathrm{L}} = 2.42$ and $\eta^{\mathrm{NL}} = 4.28$, as indicated in [98, Tab. A.2.1.1.2-3].

Finally, it is worth mentioning that the simulator is calibrated accordingly to the 3GPP reference scenarios described in [98, Annex 2.2] that use the list of parameters indicated in [98, Table A.2.2-1].

Table 3.2: 3GPP-based system-level simulation parameters

mMIMO-BS	**Description**
Cellular layout	Wrap around hexagonal, 19 sites, 3 sectors/site
Deployment	ISD: 500 m, height: 32 m
Antenna array	ULA with element spacing 0.5λ, Jakes correlation model, number of antennas per array: 64
Antenna pattern	70° H x 10° V beamwidths, 14 dBi max, downtilt: 15°
Precoder	Zero-forcing
Tx power/Noise figure	46 dBm, 5 dBm
Self-BH SC	**Description**
Deployment	Random: $\{4, 8, 16\}$ SCs/sector on average, Ad-hoc: 16 SCs/sector on average, height: 5 m
Backhaul antenna pattern	Single antenna, 5 dBi antenna gain, Omni
Access antenna pattern - Patch	Single antenna, 80° H x 80° V beamwidths, 5 dBi max, downtilt: 90°
Access antenna pattern - Yagi	Single antenna, 58° H x 47° V beamwidths, 10 dBi max, downtilt: 90°
Tx power/Noise figure	30 dBm, 5 dB
UE	**Description**
Deployment	Random, 16 UEs/sector on average all served, height: 1.5 m
UE mobility	0 m/s
Tx power/Noise figure	23 dBm, 9 dB
Channel	**Description**
Scenario	Outdoor SC, outdoor UE
Bandwidth/Time-slot	10 MHz at 2 GHz, $Q_t = 50$ RBs, $T = 1$ ms
LoS probability, pathloss and shadowing	mMIMO-BS to UE (3GPP macro to UE models [98])
	mMIMO-BS to SC (3GPP macro to Relay Node (RN) models [98])
	SC to UE (3GPP RN to UE models [98])
Fast fading	Rician, distance-dependent K factor
Thermal noise	-174 dBm/Hz power spectral density

3.5.2 Massive MIMO Self-Backhaul: Random vs Ad-hoc Small Cell Deployments

In this subsection, we analyse the 3GPP-based simulation results for the two SC topologies described in Sec. 3.2.1, namely the ad-hoc and random SC deployments. In both cases, $\mu_u = 16$ UEs are deployed per sector on average and scheduled in access time-slots by their serving SCs. We evaluate the impact of densification by considering $\mu_b = \{4, 8, 16\}$ SCs per sector on average for the case of random SC deployments. In the ad-hoc deployment, we deploy a constant number $\mu_b = 16$ of SCs per sector on average, whereas we test different 2-D distances d separating the UE from the SC. The resource partition α is set to 0.5 to distribute between backhaul and access the available resources equally.

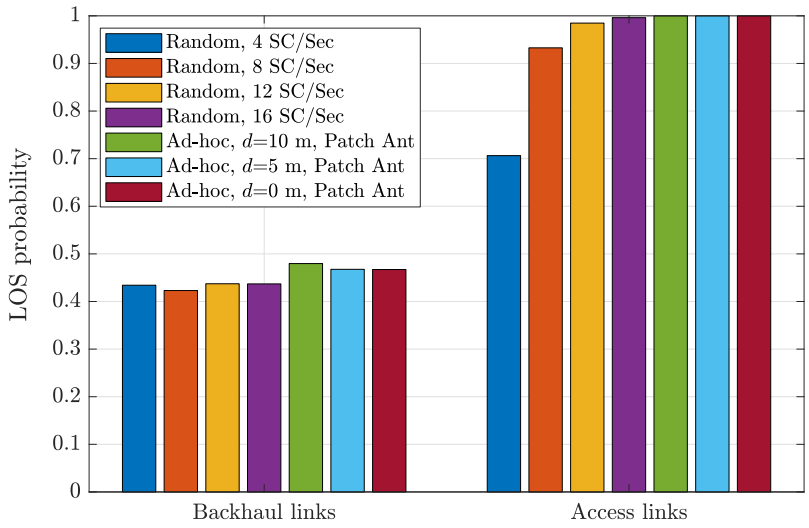
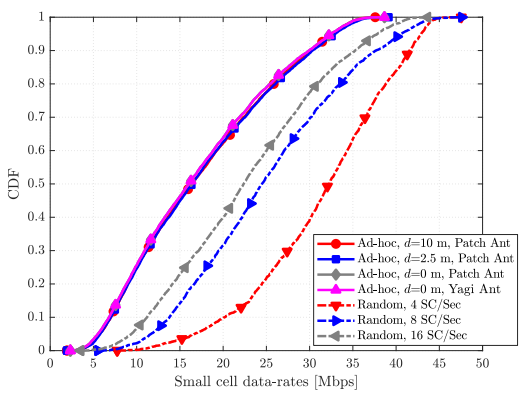


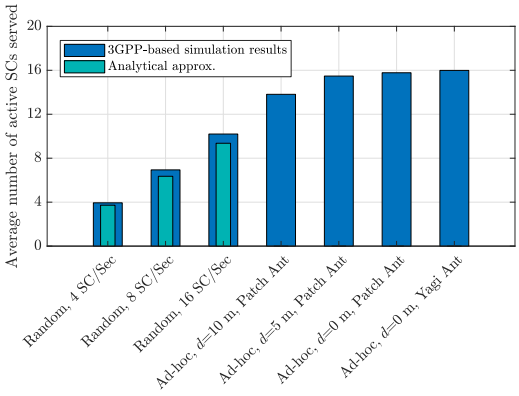
Figure 3.4: LoS probability for SC backhaul links and UE access links in s-BH networks with random and ad-hoc deployments.

As a first step, we collect the LoS probability statistics from the simulations, counting the fraction of LoS backhaul and access links in the network for different SCs densities and deployments. The links' channel is computed using tailored LoS probability functions as indicated in [98, Tab. A.2.1.1.2-3] and account for the backhaul correction factor indicated in [98, Tab. A.2.1.1.4-2]. We can summarise our analysis as follows:

- *LoS statistics of the backhaul links:* The group of results on the left part of Fig. 3.4 shows that the percentages of backhaul links in LoS are almost the same in both the random and ad-hoc deployments. In the first case, the backhaul LoS probability does not change, increasing the density of SCs. This is explained since a larger number of SCs randomly distributed in the mMIMO-BS coverage area do not change the average distance from the SCs to the mMIMO-BSs. At the same time, in the ad-hoc deployment, the SCs are positioned close to the UEs locations, which are randomly distributed in the mMIMO-BS coverage area. Therefore, the SCs LoS probability become similar to the ones of the UEs and follows the same dynamics of the random deployment. Thus, the LoS probability results of the ad-hoc deployment are similar to the ones of the random SCs deployment in the range of values of the SC-UE distances d.
- *LoS statistics of the access links:* The right part of Fig. 3.4 shows the LoS probability of the link from SC to UE in the s-BH architecture. The percentage of access links

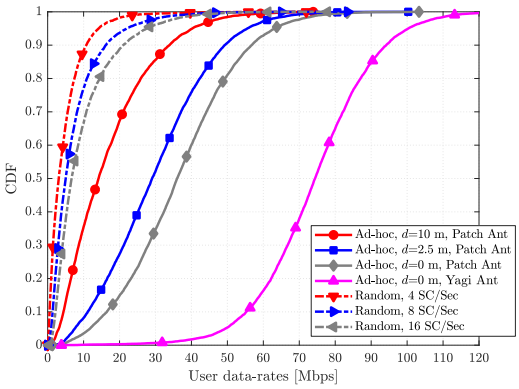


(a) CDF of SC rates for backhaul links

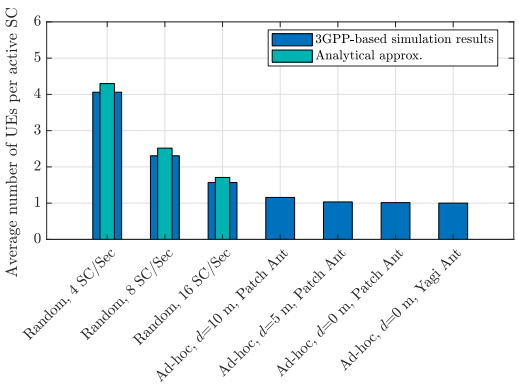


(b) Average number of active SCs served per mMIMO-BS

Figure 3.5: (a) CDF of SC rates for backhaul links and (b) the average number of active SCs served in the backhaul time-slots. (b) also shows the analytical results for the SCs activation probability in Eq. (3.10).



(a) CDF of UE rates for access links



(b) Average number of UEs served per active SC

Figure 3.6: (a) CDF of UE rates for access links, and (b) the average number of UEs served in the access time-slots. (b) also shows the analytical results of the mean number of UEs served by an active SC, given by $\lambda_u/\widetilde{\lambda}_b$.

in LoS increases as the density of SCs increases, indicating that the LoS access link probability benefits while increasing the number of SCs randomly deployed. This is caused by the shorter distance from SC to UE. Conversely, the ad-hoc deployment reaches 100% probability of LoS channel condition for all the distances d, indicating the advantage of deploying the SC targeting the UEs positions.

As a second step, in Figs. 3.5a, 3.6a and 3.7, we compute the backhaul and access data rate performance. We initially consider the two links separately and then their combined effect on the end-to-end UE rate. The following considerations can be made:

- *Backhaul links data rate performance*: Fig. 3.5a illustrates the CDFs of the backhaul

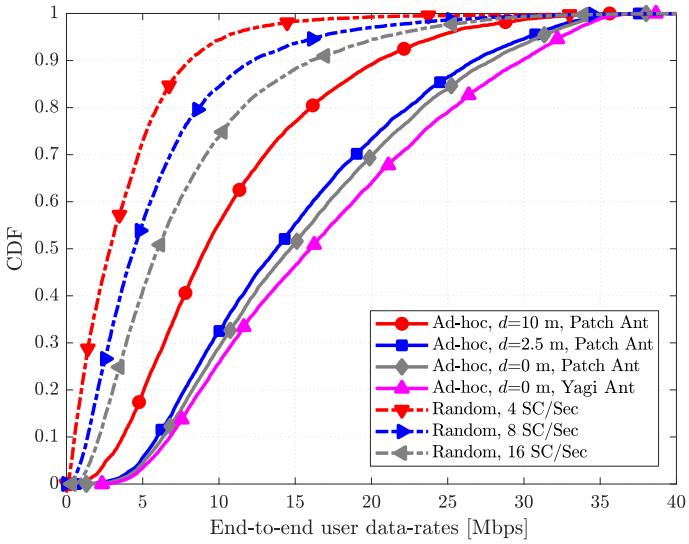


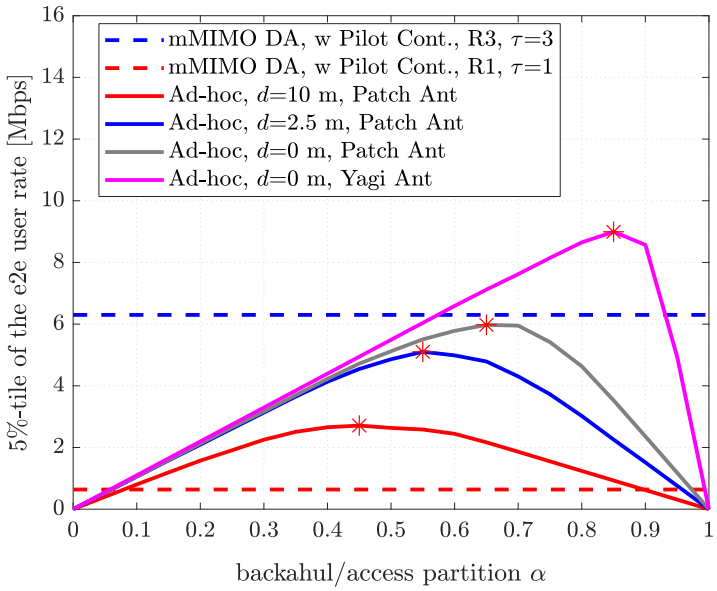
Figure 3.7: CDF of end-to-end UE rates in: (i) ad-hoc deployment of 16 SCs per sector with variable UE-to-SC distance d, and (ii) random deployment of SCs.

data rates. These results show how increasing the number of SCs randomly deployed, especially with the ad-hoc deployment, the backhaul data rate received by each SC decreases. This is due to the reduction of the multiplying factor $\left[(M-\widetilde{\mu}_b+1)P_i^{\mathrm{dl}}\right]/\widetilde{\mu}_b$ in Eq. (3.11), and the split of the transmit power among the active backhaul streams. It is worth noting that only the SCs with associated UEs are active (i.e. transmitting to the UEs) and are served via multiple backhaul links. Thus, looking at Fig. 3.5b, which shows the average number of SCs served by the mMIMO-BS when applying the random and ad-hoc SCs deployment strategies, we can better explain the results presented in Fig 3.5a. While with the ad-hoc deployment, almost all the 16 SCs are always active, with 16 randomly deployed SCs, only 10 of them are active on average, due to the UEs association procedure. As depicted in Fig. 3.5b, the analytical approximation in Eq. (3.10) of the SC activation probability matches the numerical results obtained by simulations.

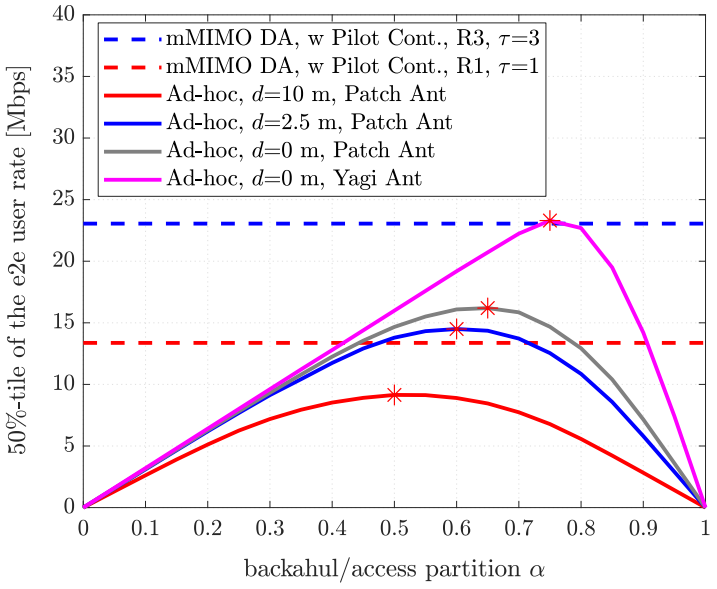
- *Access links data rate performance*: Fig. 3.6a shows the results for the access data rate. As a general conclusion, we can see that adding more randomly deployed SCs in the sector does not introduce a significant gain, while opportunistically deploying one SC closer to each UE is quite beneficial. In the following, we discuss the different factors playing a key role in these results. On the one hand, when densifying the network, the carrier signal benefits from having SCs more likely in close vicinity to

the served UE, even if a random SC deployment does not always guarantee this vicinity. Also, each SC has to serve a progressively reduced number of UEs in the access links (as indicated in Fig. 3.6b), and accordingly in the backhaul links, having more RBs available to allocate to each UE over different links. On the other hand, adding more SCs increases the probability of having a larger number of LoS interferers at the UE side. In the random deployment, the interference links' power grows faster than the carrier signal power due to NLoS to LoS transition of the interference links [30]. In the ad-hoc case, the same interference effect takes place. However, decreasing the distance d separating the UE and the SC ensures that the SC's position is always close to the served UE. Therefore, the carrier signal's power increases faster than the interference power, and thus the hit in the SINR is not as significant. As a result, only a very dense deployment of random SCs could provide the same data rate as the ad-hoc deployment. In Sec. 3.5.5, we will discuss the asymptotic behaviour when increasing the density of SCs, and quantify the number of required randomly deployed SCs to achieve this condition. Fig. 3.6a shows how equipping the SC with a more directive antenna, e.g. Yagi, compared to the Patch antenna further boosts the access link capacity to achieve 75 Mbps per UE at the median value. Two complementary effects cause this performance enhancement: *i)* the signal improvements provided by the higher antenna gain of the directive Yagi, and *ii)* the reduced interference created towards neighbour UEs served by other SCs.

- *End-to-end performance*: Fig. 3.7 shows the end-to-end results given by the combination of the two-hop, backhaul and access, performance previously depicted in Figs. 3.5a and 3.6a. Overall, the end-to-end data rates of the random deployment are more limited by the access links than by the backhaul links, as shown by comparing the results in Figs. 3.7 to the one in 3.6a. On the contrary, the end-to-end data rates of the ad-hoc deployment outperform the random deployment ones but are severely penalised by the backhaul links, as shown by comparing the results in Figs. 3.7 to the one in 3.5a. Thus, the reduced backhaul capacity of the mMIMO s-BH ad-hoc deployment does not fully exploit the potentially high data rate achieved in the access. This behaviour suggests the need to optimise the splitting of resources between the two links. Indeed, a significant improvement in end-to-end rates would be achieved by



(a) 5-th percentile



(b) 50-th percentile

Figure 3.8: (a) 5-th, and (b) 50-th percentile of the end-to-end UE rates as a function of the fraction α of backhaul time-slots.

allocating more resources to backhaul links. This is analysed in the following section.

3.5.3 Massive MIMO Self-Backhaul: Access and Backhaul Resource Allocation

In Fig. 3.8, we vary α in the range $0 \leq \alpha \leq 1$, and analyse the behaviour of UEs rate at the 5-th and 50-th percentiles of the CDF. The configurations $\alpha = 0$ and $\alpha = 1$ entail that

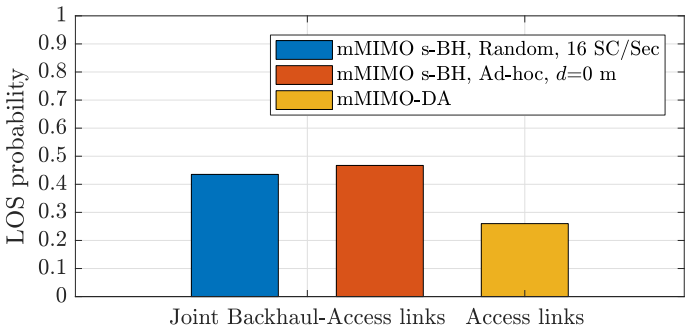


Figure 3.9: LoS probability for joint backhaul-access links in the s-BH network and access links in the DA network.

the entire time-slot duration T is allocated to the access and the backhaul, respectively. Therefore, the UE rates for these two values are equal to 0, since no resources are left for the other link. Moreover, the configuration α^* represents the value of α that maximises the UE rate. For instance, with $d = 0$ and Yagi antennas at the SCs, α^* is equal to 0.85 when looking at the 5-th percentile. Fig. 3.8 brings the following insights:

- By comparing the results between Fig. 3.8a and Fig. 3.8b, it is important to note that the optimal α changes from 0.85 to 0.75. A tradeoff exists between 5-th and 50-th percentile performance, and they cannot be optimised simultaneously. Assuming that the network uses $\alpha = 0.85$, which is the optimal value for cell edge UEs (5-th percentile of the CDF), the median UEs (50-th percentile of the CDF) can achieve an end-to-end rate of 19.5 Mbps, which decreases by 16% the maximum end-to-end rate achievable of 23.3 Mbps with $\alpha = 0.75$.
- In Figs. 3.8a and 3.8b, we depict with dashed lines the mMIMO DA network results. The figures show that a properly designed mMIMO s-BH radio resource partitioning can improve the performance of the cell edge UEs and maintain the same performance of mMIMO DA architecture for the median UEs. A more detailed comparison is further developed in the next section.

3.5.4 Massive MIMO Architectures: Self-Backhaul vs Direct Access

First, we compare the joint probability of LoS-LoS channels in the backhaul-access legs for the s-BH network with the probability of LoS channel in the access links for the DA network. As shown in Fig. 3.9, the joint backhaul-access LoS probabilities of the s-BH architectures,

i.e. 43% and 47% for random and ad-hoc, respectively, are higher than the LoS probability of the DA architecture, i.e. 25%. Those are the cases where the s-BH architecture can improve the UE performance of the DA architecture.

We also compare the end-to-end UE rates resulting from the 3GPP-based simulations of the mMIMO s-BH and mMIMO DA networks. As shown in Fig. 3.10, the mMIMO s-BH network with the ad-hoc deployment of SCs provides better performance than the mMIMO DA network with pilot reuse 1 at the bottom of the CDF, i.e. below the 50-th percentile. Pilot contamination severely degrades UEs' data rates at the cell edge in the mMIMO DA setup with pilot reuse 1. On the other hand, in the s-BH network, due to the longer coherence time of the backhaul channel T_{BH} compared to the system time-slot duration T there is no pilot contamination. As shown in Fig. 3.9, the UEs benefit from the proximity of SCs, which reduces the pathloss and improves the LoS propagation condition.

However, by adopting the pilot reuse 3 in the mMIMO DA network, the pilot contamination effect reduces, and the results show that the mMIMO DA performance exceeds the one of the mMIMO s-BH with $\alpha = 0.5$, even if the pilot overhead ($\tau = 3$ OFDM symbols) is 3 times larger than with pilot reuse 1. The mMIMO s-BH architecture provides the same performance as the mMIMO DA for the median UEs, only when the optimal partition $\alpha = \alpha^*$ is selected. Overall, mMIMO s-BH under-performs mMIMO DA above the median of the UE rates and provides rate improvements for cell edge UEs that amount to 30% and a tenfold gain adopting mMIMO DA pilot reuse 3 and reuse 1, respectively.

3.5.5 Asymptotic Data Rate Analysis

Figs. 3.11a and 3.11b show the convergence behaviour of the backhaul and access data rates for the random SCs deployment compared to the results obtained with the ad-hoc deployment of 16 SCs, positioned at fixed distance $d = 0$ from the UEs locations. In both cases, the results are obtained by numerical integration of Eqs. (3.16) and (3.19). Fig. 3.11a show the convergence results of the backhaul link. The backhaul data rate of the random distribution (solid line) converges to the one of the ad-hoc (dashed line) when the number of SCs is 100 times larger than the number of ad-hoc SCs (denoted by the marker "*"). Fig. 3.11b shows the convergence results of the access link. In this case, the data rate of the

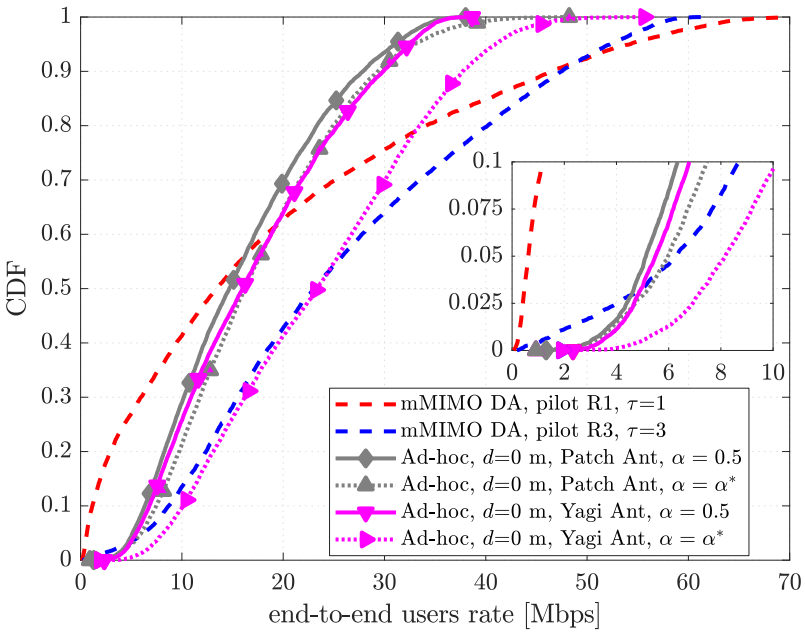


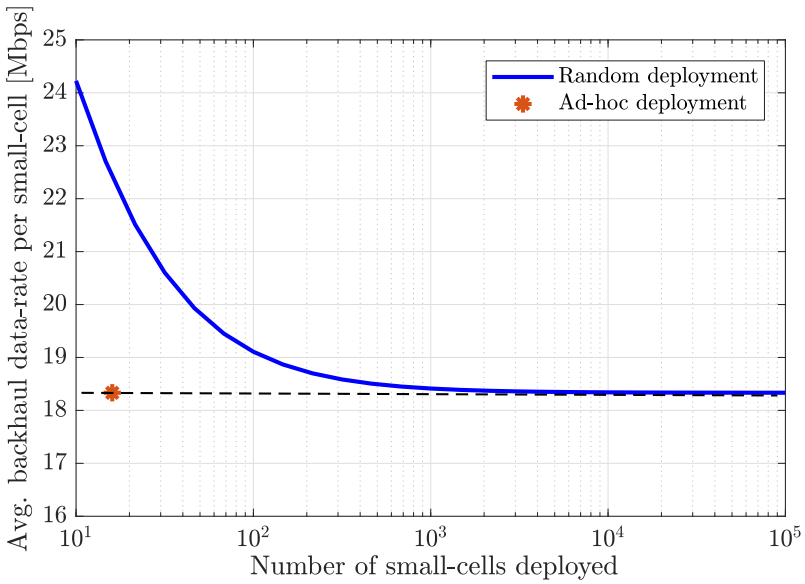
Figure 3.10: Two types of curves are represented: (*i*) mMIMO DA with pilot reuse schemes 1 and 3; (*ii*) ad-hoc deployment of 16 SCs per sector for $\alpha = 0.5$ and $\alpha = \alpha^*$. For the latter partition value, the 50-th percentile of the end-to-end UE rate is maximised (as shown in Fig. 3.8b).

random (solid line) converges to the ad-hoc (dashed line) when the number of SCs is 1000 times larger than the number of SCs deployed in the ad-hoc case (marker "*"). As shown in Sec. 3.5.3, the ad-hoc deployment is the one that maximises the end-to-end data rates of two-hop communication.

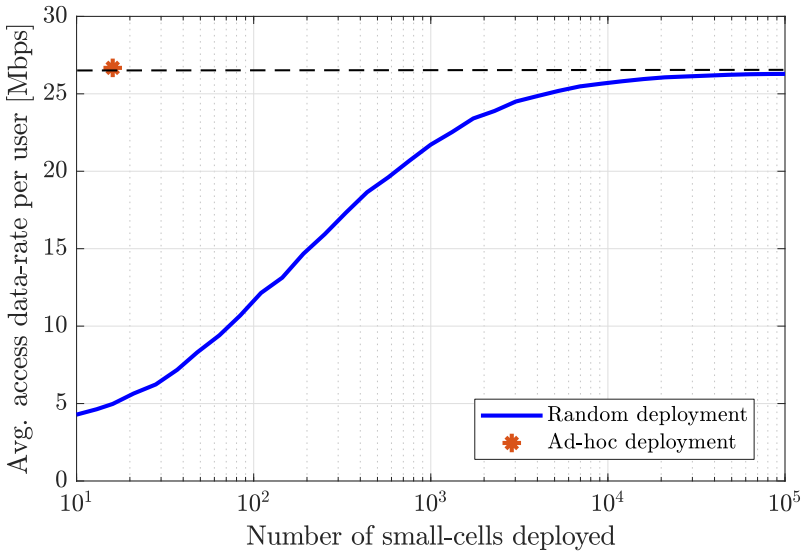
The main takeaway is summarised by the possibility for the MNOs to explore the ad-option of future dynamic SCs network infrastructures [120]. Indeed, instead of significantly over-provisioning the number of SCs, it may be beneficial to dynamically re-position only the active ones, guaranteeing the same performances obtained with a very dense deployment of SCs.

3.6 Conclusion

In this chapter, we studied the performance results for two 5G mMIMO architectures working at frequencies below 6 GHz: mMIMO s-BH and mMIMO DA. In the mMIMO s-BH architecture, we analysed two configurations: random deployment of SCs in the coverage area of the serving macro cell and ad-hoc deployment of SCs close to each UE. This study's first takeaway is that the random SCs distribution entails deploying thousands of SCs to



(a) Asymptotic SC data rate for backhaul links



(b) Asymptotic UE data rate for access links

Figure 3.11: Asymptotic performance measures for backhaul and access links in the s-BH network when random and ad-hoc deployments of SCs are considered.

achieve the upper bound access link performance. On the other hand, the ad-hoc deployment benefits from the proximity of the SCs to the UEs and outperforms the random one for reasonable numbers of deployed SCs. However, the SC requires knowing the UE position, which is complicated to track in case of UE mobility. The second takeaway is that an ad-hoc SCs deployment supported by mMIMO s-BH provides rate improvements for cell edge UEs that amount to 30% and a tenfold gain compared to mMIMO DA with pilot reuse 3 and reuse 1, respectively. This highlight the detrimental effect of the pilot contamination, which

severely impacts the UEs performance at the cell edge where it is better to serve the UEs with the mMIMO s-BH architecture. On the other hand, mMIMO DA outperforms s-BH above the median of the UE rates, meaning that when pilot contamination is less severe and the LoS probability of the DA links improves, is better to serve the UEs with the mMIMO DA architecture.

Future works will focus on improving the backhaul links' capacity, which is currently the main limitation in the s-BH architecture at the Sub-6 GHz frequencies. This improvement can be realised using the millimetre Wave (mmWave) frequencies, which can guarantee the availability of abundant spectrum in the mmWave bands that enable high data rate. The next chapter of this dissertation will address this specific aspect.

4 Analysis of mmWave Indoor Network Performance with Detection of Dynamic Blockers

In this chapter, we analyse a millimetre Wave (mmWave) indoor network scenario affected by dynamic blockages moving in the environment. We introduce a method inspired by the Third Generation Partnership Project (3GPP) beam failure recovery (BFR) procedure that considers the real-time detection of blockages to limit the temporary data rate losses and improve the air interface latency during the blocked time instants. The technical work of this chapter is based on the work presented in the first part of our paper [21].

4.1 Introduction

Millimetre-wave communication is one of the key technologies for deploying the 5-th Generation (5G) and beyond Radio Access Networks (RANs) due to its wide spectrum bandwidth, which can potentially support Enhanced Mobile BroadBand (eMBB) services and Ultra-Reliable Low-Latency Communication (URLLC) [4, 10, 52, 121]. MmWave cellular communications are well-suited for indoor network deployments, e.g., open offices, shopping centres and industrial spaces [71]. Adopting multi-antenna arrays with analog-only or hybrid architectures enables to form narrow beams that provide high directivity gains and compensate for the severe pathloss [53]. At the same time, mmWave communication is very sensitive to blockage events [59, 60, 122]. For example, these occur when dynamic objects such as humans, robots or vehicles, which move in the environment, obstruct the signal path causing the channel propagation conditions to change from Line-of-Sight (LoS) to Non

Line-of-Sight (NLoS) rapidly, impacting the data rate and latency performance [51, 78].

New applications, like the control of a programmable logic controller (PLC) in automated apparatus, demand end-to-end latency in the order of milliseconds [70]. Such a latency cyclic time requires the air-interface Physical layer (PHY) transmission time to be approximately 100 μs per direction, while leaving the remaining time budget to the other system operations, such as the embedded device processing and application processing [69]. Moreover, the analog beamforming architecture for mmWave systems enables to form high directional beams but is limited to point towards one beam direction at a time. Consequently, only User Equipments (UEs) that are associated to the same beam can share the spectrum resources between multiple UEs simultaneously, like in the Frequency-division multiple access (FDMA) adopted in Long Term Evolution (LTE). On the other hand, UEs associated with different beams are served in time one after the other, adopting the Time-division multiple access (TDMA) scheme. Given the utilisation of TDMA in mmWave systems and the sensitivity of mmWave communication to blockage, there are many interrogations about whether the mmWave can support low-latency requirements. Several solutions have been considered in New Radio (NR) system design to support low-latency data transmissions [51]. We briefly explain three of them as follows:

- *Scalable Transmission Time Interval (TTI):* Thanks to the wider coherence bandwidth at mmWave frequencies, the NR Standard considers operations in the frequency range 2 (FR2) with larger sub-carrier spacing than the Sub-6 GHz frequency range 1 (FR1). This enables to reduce the Orthogonal Frequency Division Multiplexing (OFDM) symbol duration to 8.33 μs and 4.16 μs for the sub-carrier spacings $\Delta_f = 120$ kHz and $\Delta_f = 240$ kHz, respectively [123]. As a consequence, the TTI duration reduces to 0.125 msec and 0.0625 msec for the OFDM symbol duration of 8.33 μs and 4.16 μs, respectively.
- *Mini-slot:* To facilitate the transmission of small packets and control messages requiring low latency, NR Standard introduces the possibility to schedule the transmission in a mini-slot, which occupies only a small part of the TTI, e.g. 2, 4 or 7 OFDM symbols over the 14 OFDM symbols available in the TTI, and it can be scheduled starting from an arbitrary OFDM symbol within the TTI [124].

- *Digital beamforming for control:* The analog beamforming architecture represents a major drawback for transmitting shorts packets containing control information since with beamforming in place the Base Station (BS) is limited to transmit the control data towards one UE at a time. To reduce the overhead to transmit the control information to multiple UEs, the method described in [125] proposes to switch to a digital beamforming architecture only during the control transmission phase and simultaneously transmit control data to several UEs. Then, during the data transmission, the analog beamforming architecture is used to deliver high beamforming gain and provide high signal-to-noise ratio (SNR) that support the transmission with higher-order Modulation and Coding Scheme (MCS) and packets of large sizes.

One major challenge remains to guarantee a high data rate over the PHY interface for the whole packet transmission duration, even if the blockage affects the link quality. This a major impediment for the deployment of mmWave networks and may limit the utilisation of this technology for the most compelling low-latency use cases, e.g. closed-loop control in process automation and mobile robots [70].

4.1.1 Scope and Contributions

In this chapter, we analyse the data rate and latency performance of a mmWave network with dynamic blockers moving in the environment and obstructing the beam pair used to connect the UE. We consider a mmWave Indoor cellular network where each mmWave Base Station (mmWave-BS) employs a 2D beamforming codebook containing multiple beamforming vectors covering different parts of the mmWave-BS sector area. To contrast the significant data rate losses occurring during the blocked time instants, we introduce a radio link control method inspired by the beam failure recovery (BFR) procedure adopted in the 3GPP NR Standard [86]. This method detects the blockage events when the received signal power drops below a given threshold. At this point, the detection triggers the beam recovery mechanism that searches for an alternative beam to re-configure the transmission. Our contributions in this chapter can be summarised as follows:

1) We analyse the mmWave Direct Access (DA) communication in an indoor network scenario affected by dynamic blockages through system-level simulations. We adopt

a geometric blockage model to gain more insights about the stability of the UEs data rate under the temporal transition from LoS to NLoS channel propagation conditions due to the blockage effect.

2) We introduce a beam recovery method in the mmWave network operations that performs real-time detection of blockage events and recovers the mmWave connections by switching to an alternative beam in case of blockage.
3) We provide a detailed analysis of the UEs performance testing key aspects that characterise the dynamic blockage effect, such as the blocker speed and the blockage probability. This analysis highlighted the impact of the dynamic blockers on the data rate and the probability that the air interface latency complies with a realistic value of the PHY transmission duration.

The remainder of this chapter is organised as follows. Section 4.2 describes the system model and presents the beam recovery based on the detection method; Section 4.3 presents the simulations and numerical results of the mmWave indoor network adopting the detection method and Section 4.4 summarises the key findings.

The following notation is used throughout the chapter: boldface lower case and boldface upper case are used for column vectors $\mathbf{x}$ and matrices $\mathbf{X}$. $\mathbf{x}^{\mathsf{T}}$ denote the transpose of $\mathbf{x}$. A Gaussian random variable x is denoted $x \sim \mathcal{CN}(\mu, \sigma^2)$, where μ is the mean and σ^2 is the variance.

4.2 System Model

4.2.1 Network Layout

We considered the network layout depicted in Fig. 4.1 and inspired by the 3GPP Indoor Hot-spot scenario [57, 126], formed by four cell sites deployed on a rectangular grid with an Inter Site Distance (ISD) of 20 meters. Each cell site has three sectors, served respectively by three mmWave-BSs oriented with an angle θ_j with $j \in \{1, 2, 3\}$ and operating at mmWave carrier frequency f_c, with bandwidth BW. We assume a set of $\mathcal{K}$ UEs distributed uniformly at random over a rectangular area of 50 meters by 40 meters. We consider dynamic blockers

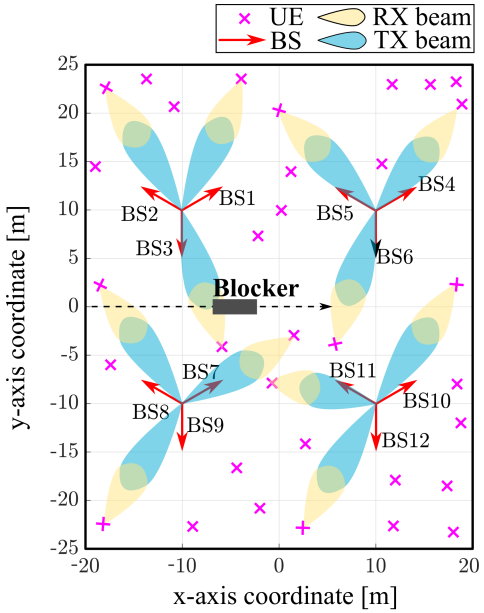


Figure 4.1: Layout of the indoor mmWave-BS network.

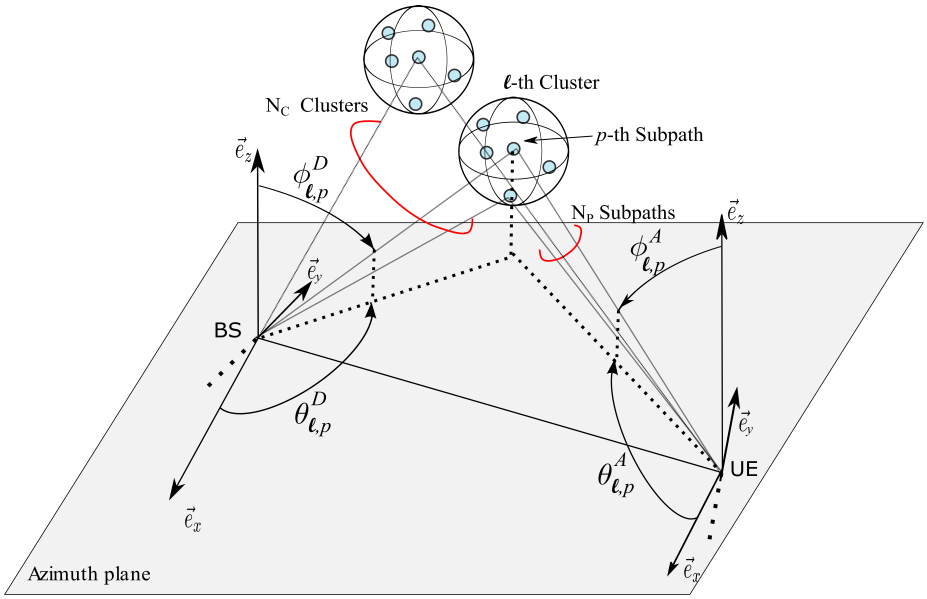


Figure 4.2: Representation of the 3D channel with multiple clusters and sub-paths.

having dimension $w \times h$ moving at a constant speed v following a linear trajectory at the centre of the scenario along the x-axis.

4.2.2 3D Channel Model

We generate the mmWave channel between a BS and UE following the Three-Dimensional (3D) Spatial Channel Model (SCM) specified in the 3GPP Technical Report (TR) 38.901

[57]. In this document, 3GPP describes a geometric-based stochastic channel model that accounts for a scattering environment formed with N_C clusters that are in turn composed by N_P sub-paths, as depicted in Fig. 4.2. The channel impulse response between the Transmitter (TX)-Receiver (RX) antenna pair can be expressed as [57, 127, 128]

$$\begin{aligned} H_{u,s}(t,\tau) = \sqrt{10^{-\frac{PL+\sigma_{SF}}{10}}} \sum_{\ell=1}^{N_C} \sum_{p=1}^{N_P} & \mathbf{F}_{rx}(\theta^A_{\ell,p}, \phi^A_{\ell,p}) e^{j\Phi_{\ell,p}} \mathbf{F}_{tx}(\theta^D_{\ell,p}, \phi^D_{\ell,p}) \\ & \times e^{j\left(\mathbf{k}^{\mathsf{T}}_{rx,\ell,p} \cdot \mathbf{d}_{rx,u} + \mathbf{k}^{\mathsf{T}}_{tx,\ell,p} \cdot \mathbf{d}_{tx,s}\right)} \\ & \times \sqrt{P_{\ell,p}} \cdot 10^{-\frac{BL_{\ell,p}(t)}{20}} \delta(\tau - \tau_{\ell,p}), \end{aligned} \tag{4.1}$$

where $\sqrt{10^{-\frac{PL+\sigma_{SF}}{10}}}$ represents the slow channel gain, which includes pathloss PL and shadowing σ_{SF}. For each sub-path p in cluster ℓ, the model specifies Angle of Arrival (AoA) $(\theta^A_{\ell,p}, \phi^A_{\ell,p})$ and Angle of Departure (AoD) $(\theta^D_{\ell,p}, \phi^D_{\ell,p})$, which modify the RX and TX antenna field patterns $\mathbf{F}_{rx}$ and $\mathbf{F}_{tx}$, and $\Phi_{\ell,p}$, which represents a random initial phase if polarisation is not considered.

Moreover, the terms $\exp(j\mathbf{k}^{\mathsf{T}}_{rx,\ell,p} \cdot \mathbf{d}_{rx,u})$ and $\exp(j\mathbf{k}^{\mathsf{T}}_{tx,\ell,p} \cdot \mathbf{d}_{tx,s})$ represent the array responses of the receiving and transmitting antennas, where $\mathbf{k}_{rx,\ell,p}$ and $\mathbf{k}_{tx,\ell,p}$ are the RX and TX wave vectors along to the directions of the p-th sub-path in cluster ℓ such that $\|\mathbf{k}\| = \frac{2\pi}{\lambda_0}$. Additionally, $\mathbf{d}_{rx,u}$ is the location vector of the receiving antenna s whereas $\mathbf{d}_{tx,s}$ is that of the transmitting antenna u computed in the global Cartesian coordinate system $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$ (see Fig. 4.2). Finally, for each sub-path p in cluster ℓ, the model accounts for the power gain $P_{\ell,p}$, blockage loss $BL_{\ell,p}(t)$ and propagation delay $\tau_{\ell,p}$.

To note that Eq. (4.1) represents the channel impulse response for the NLoS case. The channel impulse response for the LoS case can be obtained by adding the LoS channel coefficient to Eq. (4.1) and scaling both terms by the Rician K-factor K_R as

$$H^{\text{LoS}}_{u,s}(t,\tau) = \sqrt{\frac{1}{K_R+1}} H^{\text{NLoS}}_{u,s}(t,\tau) + \sqrt{\frac{K_R}{K_R+1}} H^{\text{LoS}}_{u,s,1}(t)\delta(\tau - \tau_1), \tag{4.2}$$

where $H^{\text{LoS}}_{u,s,1}(t)$ and τ_1 are the channel impulse response and the propagation delay of the LoS path, respectively [127].

4.2.3 Blockage Model

To model the blockage effect we consider the 3GPP Blockage Model B, which introduces a time-dependent component $BL(t)$ that reduces the power of the 3-D channel cluster's sub-path when intersecting the blocker [57]. The Blockage Model B deploys a 3D rectangular screen with sizes $w \times h$ for each blocker moving in the scenario and computes the total power loss given by the knife-edge diffraction from the edges of the screen [129]. Diffraction makes the electromagnetic (radio) waves to bend around the screen's corners in their path, spreading to the shadow region behind the screen. The RX in the shadow region is never totally blocked. Nevertheless, only part of signal power diffracted by the blocker's edges reaches the RX position. Therefore, the total diffraction loss accounts for the signal power difference between the signal diffracted by the blocker's edges and the signal propagated under free-space propagation conditions, i.e. without the screen.

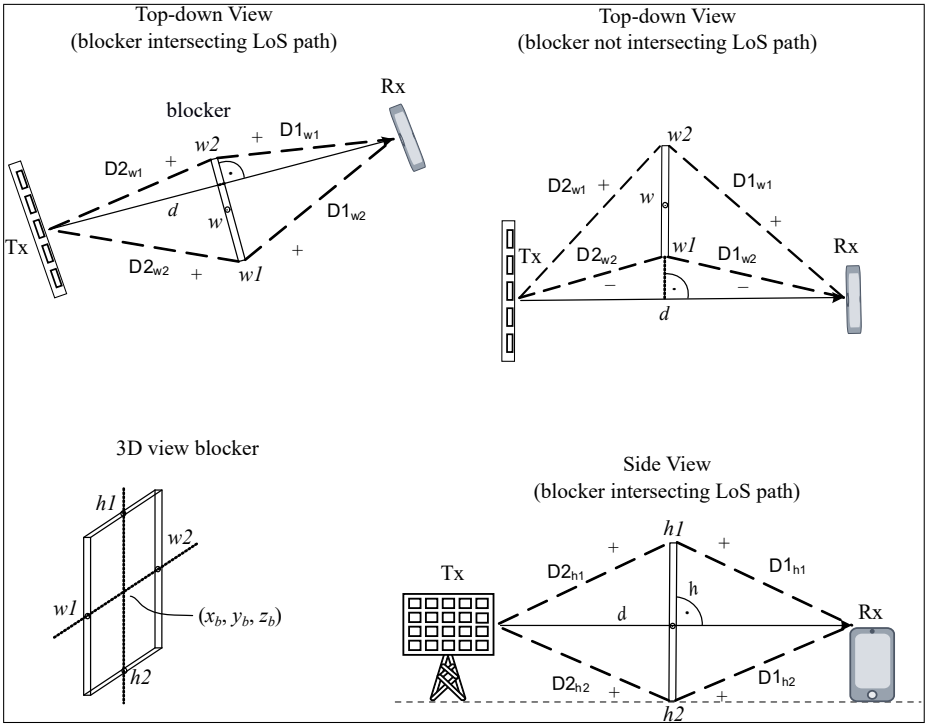
The Blockage Model B maintains consistent results over time, space, and channel frequency components because of its geometric approach. This aspect is especially crucial to model directional mmWave transmission if UEs are positioned close to each other. In fact, when the UEs direct propagation paths towards the BS are closely spaced in the angular domain, the blocker intersects their paths in close proximity, causing a blockage loss that follows similar temporal dynamics, and requires to be consistent for different UEs.

Moreover, the blocker may obstruct the LoS path from BS to UE or one or more of the NLoS paths, resulting from the reflection with static objects such as buildings, walls, etc. As illustrated in Figs. 4.3a and 4.3b, the Blockage Model B discriminates between LoS and NLoS channel propagation conditions. In the first case, the blocker intersects the direct path connecting the BS to the UE, whereas in the latter, the blocker intersects the path reflected by the static object and directed at the UE.

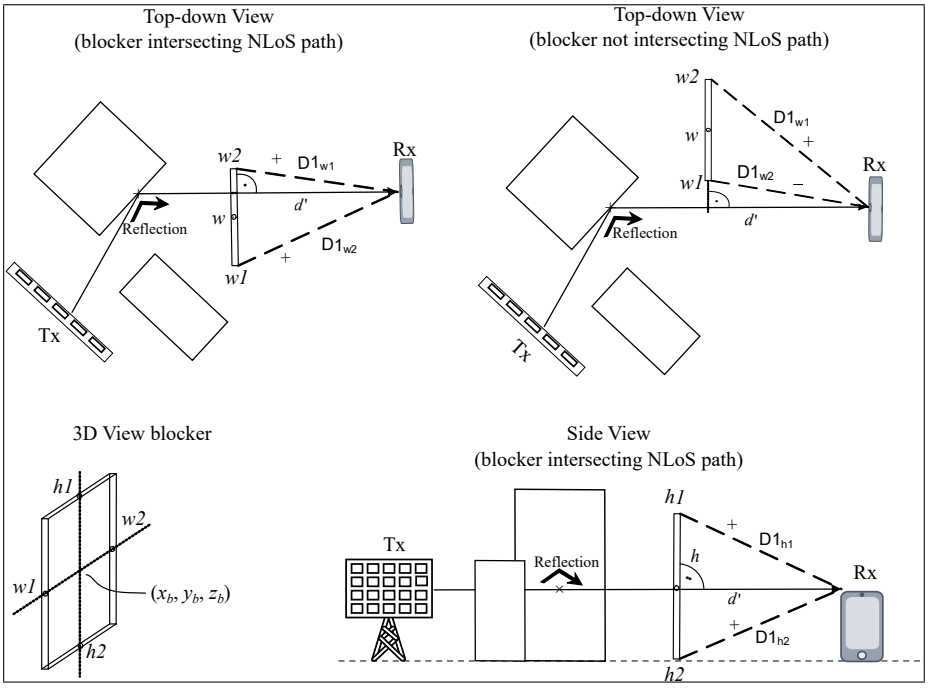
The overall blockage attenuation can be expressed as [57]

$$BL = -20 \log_{10} \left(1 - (F_{h1} + F_{h2})(F_{w1} + F_{w2})\right), \tag{4.3}$$

where $F_{h1}, F_{h2}, F_{w1}, F_{w2}$ represent the diffraction losses observed at the RX corresponding



(a) Illustration of the top-down views, 3D view and side view reporting the distances TX-blocker and blocker-RX for the LoS case.



(b) Illustration of the top-down views, 3D view and side view reporting the distances blocker-RX for the NLoS case.

Figure 4.3: Applications of the Blockage Model B for the LoS and NLoS channel propagation conditions. For both cases, we illustrate the geometric relation between blocker, RX and TX. We show the distances $D1_{w1|w2|h1|h2}$ and $D2_{w1|w2|h1|h2}$ used to compute the diffraction losses from blockers' horizontal edges defined in Eq. (4.4) and the diffraction losses from blockers' vertical edges defined in Eq. (4.5).

to the four blocker edges. For the blocker's horizontal edges, F_{w1}, F_{w2} are expressed as [57]

$$F_{w1|w2} = \begin{cases} \dfrac{\tan^{-1}\left(\pm\frac{\pi}{2}\sqrt{\frac{\pi}{\lambda_0}\left(D1_{w1|w2} + D2_{w1|w2} - d\right)}\right)}{\pi} & \text{for LoS path,} \\ \dfrac{\tan^{-1}\left(\pm\frac{\pi}{2}\sqrt{\frac{\pi}{\lambda_0}\left(D1_{w1|w2} - d'\right)}\right)}{\pi} & \text{for NLoS path,} \end{cases} \tag{4.4}$$

where λ_0 is the wavelength, d and d' are the BS-UE and blocker-UE distances, corresponding to the LoS and NLoS paths, respectively. $D1_{w1|w2}$ and $D2_{w1|w2}$ denote the blocker-RX and TX-blocker distances evaluated in the top-down view and are defined as the line segments connecting the horizontal edges of the blocker to the TX and RX antenna positions, as depicted in the top-down views of Figs. 4.3a and 4.3b. The $\pm$ sign changes whether the blocker intersects the LoS and NLoS paths. To note that if the blocker intersects the LoS and NLoS paths, the $+$ sign is applied at both edges. Differently, if the blocker does not intersect the paths, but one of the edges is still in a shadow region, the $-$ sign is applied to the edge closest to the path directed to the RX and the $+$ sign is applied to the edge farthest from the path. Conversely, for the blocker's vertical edges, F_{h1}, F_{h2} are expressed as [57]

$$F_{h1|h2} = \begin{cases} \dfrac{\tan^{-1}\left(\pm\frac{\pi}{2}\sqrt{\frac{\pi}{\lambda_0}\left(D1_{h1|h2} + D2_{h1|h2} - d\right)}\right)}{\pi} & \text{for LoS path,} \\ \dfrac{\tan^{-1}\left(\pm\frac{\pi}{2}\sqrt{\frac{\pi}{\lambda_0}\left(D1_{h1|h2} - d'\right)}\right)}{\pi} & \text{for NLoS path,} \end{cases} \tag{4.5}$$

where $D1_{h1|h2}$ and $D2_{h1|h2}$ denote the blocker-RX and TX-blocker distances evaluated in the side view and are defined as the line segments connecting the vertical edges of the blocker to the TX or RX antenna positions, as depicted in the bottom parts of Figs. 4.3a and 4.3b. We consider the blocker positioned on the ground for the blocker's vertical edges and height sufficient to intersect the LoS and NLoS paths.

Fig. 4.4 represents the blockage loss computed with Eq. (4.3). It is worth remarking that the diffraction effect is frequency-dependent, and these results are obtained for the carrier frequency $f_c = 28$ GHz; however, the blockage loss increases at higher carrier frequencies. Additionally, these results are obtained while varying the distance between blocker and RX and the blocker width, which allow us to observe different behaviours. Firstly, moving on

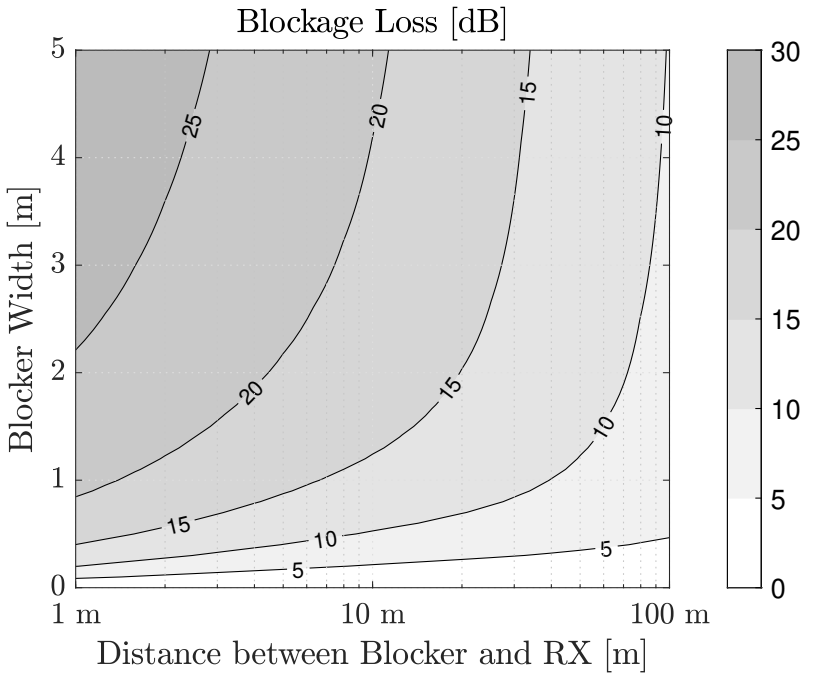
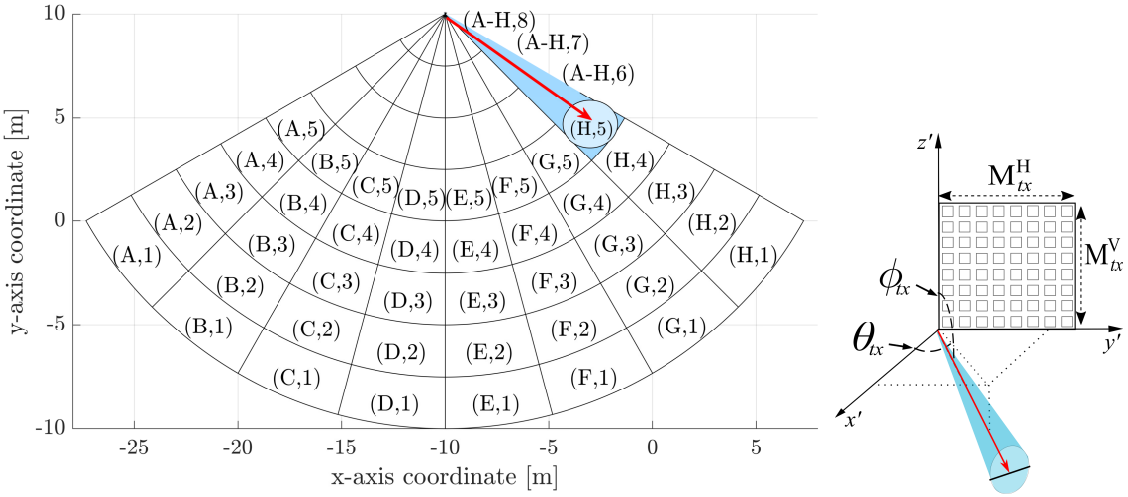


Figure 4.4: Illustration of blockage loss BL as a function of the distance between the blocker and RX (TX-blocker distance fixed to 10 m) and variable blocker width (blocker height set to 3 m) for the carrier frequency $f_c = 28$ GHz obtained with the Blockage Model B described in Sec. 4.2.3.

the y-axis direction, the blocker's width increases, whereas the blocker-RX distance remains the same. The larger blocker width implies that the shadow region behind the blocker enlarges. Consequently, the sum of the diffraction loss terms, i.e. $F_{w1} + F_{w2}$, increases to the maximum value of 1, which corresponds to heavy shadowing conditions and high blockage loss. On the other hand, moving on the x-axis direction at the fixed blocker's width, e.g. 2 m, increases the distances between the blocker and RX. The higher blocker-RX distance implies that the RX moves far away from the blocker's shadow region. Consequently, the sum of the diffraction loss terms decreases to the minimum value of 0, which corresponds to the minimum blockage loss.

4.2.4 Analog Beamforming

Both the BS and the UE employ phased antenna arrays with analog beamforming, characterised by a single Radio Frequency (RF) chain, to enable to form beams in predefined directions according to the beamforming vectors $\mathbf{b}_{tx}$ and $\mathbf{b}_{rx}$. Both BSs and UEs are equipped with Uniform Planar Array (UPA) with antenna spacing of a half-wavelength and a number of antennas $M_{tx} = M_{tx}^{\mathrm{V}} \times M_{tx}^{\mathrm{H}}$ and $M_{rx} = M_{rx}^{\mathrm{V}} \times M_{rx}^{\mathrm{H}}$ where the subscripts V and



(a) Top-down view

(b) Structure of the Tx UPA

Figure 4.5: (a) Top-down view of the area in the x-y plane illuminated by the Tx beams of the BS-3. (b) Structure of the UPA formed by $M^{\mathrm{H}}_{Tx} \times M^{\mathrm{V}}_{Tx}$ antennas and placed in the local coordinates system (x', y', z'). The Tx beam has azimuth θ_{Tx} defined between the x' axis and the Tx beam projection on the x'-y' plane and elevation ϕ_{Tx} defined between the z' axis and the Tx beam direction.

H denote the vertical and horizontal dimensions of the UPAs. We also denote as (θ_{tx}, ϕ_{tx}) the azimuth and elevation AoDs and as (θ_{rx}, ϕ_{rx}) the azimuth and elevation AoAs of the TX and RX beamforming vectors, respectively.

The TX beamforming vector used to steer the beam depends by the azimuth-elevation angles and for UPA can be expressed as

$$\mathbf{b}_{tx}(\theta_{tx}, \phi_{tx}) = \frac{1}{\sqrt{M^{\mathrm{H}}_{tx} M^{\mathrm{V}}_{tx}}} \Big[1, \ldots, e^{-j\pi(u \sin(\phi_{tx}) \sin(\theta_{tx}) + u' \cos(\phi_{tx}))}, \ldots \\ \ldots, e^{-j\pi((M^{\mathrm{H}}_{tx}-1) \sin(\phi_{tx}) \sin(\theta_{tx}) + (M^{\mathrm{V}}_{tx}-1) \cos(\phi_{tx}))}\Big]^{\mathsf{T}}. \tag{4.6}$$

Conversely, the RX beamforming vectors for UPA can be expressed as

$$\mathbf{b}_{rx}(\theta_{rx}, \phi_{rx}) = \frac{1}{\sqrt{M^{\mathrm{H}}_{rx} M^{\mathrm{V}}_{rx}}} \Big[1, \ldots, e^{-j\pi(u \sin(\phi_{rx}) \sin(\theta_{rx}) + u' \cos(\phi_{rx}))}, \ldots \\ \ldots, e^{-j\pi((M^{\mathrm{H}}_{rx}-1) \sin(\phi_{rx}) \sin(\theta_{rx}) + (M^{\mathrm{V}}_{rx}-1) \cos(\phi_{rx}))}\Big]^{\mathsf{T}}. \tag{4.7}$$

The beamforming vectors $\mathbf{b}_{tx}$ and $\mathbf{b}_{rx}$ are identified with two beam IDs l and q defined as $l = \{1, \ldots, N_{\mathrm{CB},Tx}\}$ and $q = \{1, \ldots, N_{\mathrm{CB},Rx}\}$, where $N_{\mathrm{CB},Tx}$ and $N_{\mathrm{CB},Rx}$ denote the cardinalities of TX and RX codebooks $\mathcal{B}_{tx}$ and $\mathcal{B}_{rx}$, respectively. Fig. 4.5a shows the top-down view of the area illuminated by the TX beams of the BS-3. Each TX beam of the

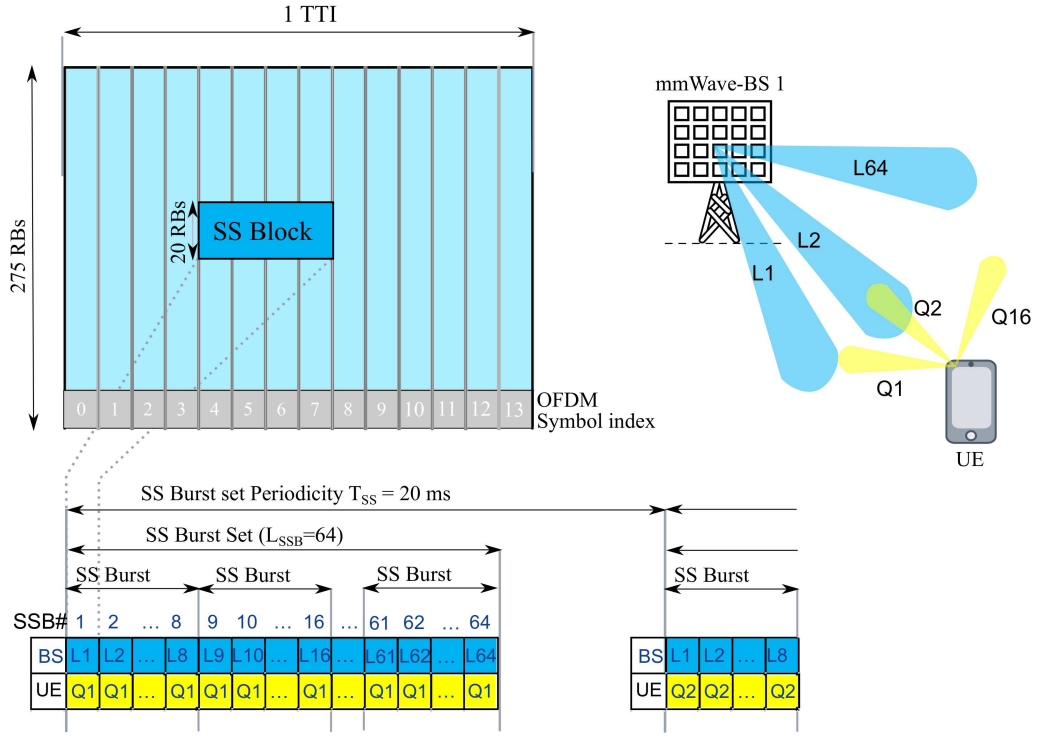


Figure 4.6: Example of structures for the synchronisation signal block (SSB) and SS Burst Set in NR Standard and related beam sweeping operations at the mmWave-BS and UE. Each SSB corresponds to a pair of TX and RX beams with identifiers (IDs) l and q, defined as $l = \{L_1, L_2, \ldots, L_{N_{\mathrm{CB},tx}}\}$ and as $q = \{Q_1, Q_2, \ldots, Q_{N_{\mathrm{CB},rx}}\}$, respectively.

codebook is steered towards the direction with AoD (θ_{tx}, ϕ_{tx}) as shown in Fig. 4.5b, and illuminates a specific area marked with a letter-number combination.

4.2.5 Downlink Data Transmission

Consider the Downlink (DL) of a NR-based system for mmWave cellular communications. The NR Standard numerology employs a physical time-frequency resource division corresponding to 14 OFDM symbols for one TTI in the time domain and 12 consecutive subcarriers in the frequency domain forming the Resource Block (RB) [123]. At a given time instant t, the DL signal transmitted from the BS j and received at the UE k can be expressed as

$$y_{j,k}(t) = \sqrt{P_b} s_k(t) \mathbf{b}_{rx}^{\mathsf{T}} \mathbf{H}_{j,k}(t) \mathbf{b}_{tx} + \mathbf{b}_{rx}^{\mathsf{T}} z_k(t), \tag{4.8}$$

where P_b is the total power of the BS, $s_k(t)$ with $\mathbb{E}[|s_k(t)|^2] = 1$ is the signal transmitted, $\mathbf{H}_{j,k}(t)$ represents the impulse response of the 3D channel, which includes the blockage effect $BL(t)$, $z_k(t) \sim \mathcal{CN}(0, \sigma_z^2)$ is the noise seen at the k-th receiver and $\mathbf{b}_{tx}$ and $\mathbf{b}_{rx}$ are the TX

and RX beamforming vectors forming the beam pair.

In the NR Standard, the beam alignment is performed through the dual TX and RX beams sweeping procedure [130]. During this procedure, the BS and UE exhaustively search among all the combinations of TX and RX beams to identify the best beam pair, referred to as primary beam pair, which provides the maximum Reference Signal Received Power (RSRP). This operation is performed during the Initial Access (IA) phase when the BSs periodically broadcast SSBs employing beamforming to improve coverage. As shown on the left side of Fig. 4.6 each SSBs is mapped to 4 OFDM symbols of the TTI in the time domain and 20 RBs over 275 RBs in the frequency domain [94]. Multiple SSBs are grouped in a SS Burst and cover successive TTIs. Multiple SS Bursts are referred to as SS Burst Set, which is transmitted in the first half-frame (5 ms) and have a periodicity of $T_{SS} = 20$ ms, corresponding to two NR frames. The maximum number of SSBs within each SS Burst Set is frequency-dependent and equals to $L_{\text{SSB}} = 64$ for the FR2 [131].

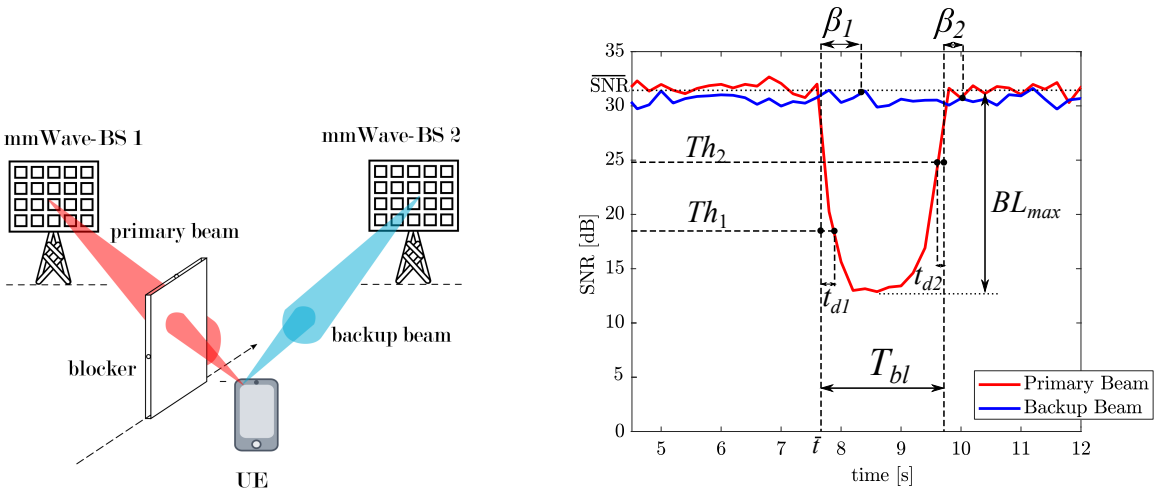
The right side of Fig. 4.6 represents the BS and UE beam sweeping procedure occurring during the SSBs transmission. The BS sweeps through its TX beam set changing the TX beam spatial direction from one SSB to another. At the same time, the UE keeps the spatial direction of the RX beam fixed for a certain number of SSBs and measures the quality of the beam pair through RSRP. The NR beam sweeping procedure duration depends on the TX and RX codebook dimensions and can take a long time to identify the pair of beams that maximises the RSRP. The time for sweeping across all the possible TX-RX combinations can be expressed as

$$T_{sweep} = T_{SS} \frac{N_{\text{CB},tx} N_{\text{CB},rx}}{L_{\text{SSB}}} + T_{SS}/2, \tag{4.9}$$

where $T_{SS}/2$ is the average time to wait until the next SS Burst Set transmission assuming that the blockage happens at a time $\bar{t} \sim \mathcal{U}[t, t + T_{SS}]$.

During the data transmission, the system operates with the Time Division Duplex (TDD) scheme, and each BS uses beam sweeping to time-multiplex multiple UEs serving one UE at a time. The data rate of the DL transmission between BS j and UE k can be expressed as

$$r'_{j,k}(t) = \frac{BW}{K_j} \log_2 \left(1 + \frac{P_b \left| \mathbf{b}_{rx}^{\mathsf{T}} \mathbf{H}_{j,k}(t) \mathbf{b}_{tx} \right|^2}{I_{j,k}(t) + \sigma_z^2} \right), \tag{4.10}$$



(a) Primary and backup beams for one of the network's UE.

(b) Variations of the SNR values for primary and backup beams during the blockage event.

Figure 4.7: Example of BR-Det method operations for one UE. At the time $t = \bar{t}$, a blocker intersects the primary beam pair. The BR-Det method detects the blockage at time $t = \bar{t} + t_{d1}$ when the SNR is lower than the threshold Th_1 and switches to the backup beam pair with a delay β_1. At the time $t = \bar{t} + T_{bl} + t_{d2}$, the BR-Det method detects that the SNR is higher than the threshold Th_2 and switches back to the primary beam pair after the blockage ends with a delay β_2.

where K_j represents the set of UEs served by the BS j and $I_{j,k}(t)$ represents the inter-cell interference. This term can be further expressed as $I_{j,k}(t) = \sum_{j' \in \mathcal{J} \setminus j} \sqrt{P_b} s_{k'}(t) \mathbf{b}_{rx}^{\mathsf{T}} \mathbf{H}_{j,k}(t) \widehat{\mathbf{b}}_{tx}$, where $\widehat{\mathbf{b}}_{tx}$ represents the TX beam directed from BS j' to UE $k' \neq k$.

4.2.6 Beam Recovery Based on Detection

In this section, we introduce a method based on a detection threshold to recover the beam when the mmWave links are affected by blockage events. In what follows, we consider that the screen blocks the TX-RX beam pair whenever it intersects the TX beam. Therefore, we will avoid the TX-RX beams distinction, and we will refer to the TX beam as a beam.

The beam recovery based on the detection (BR-Det) method operates in three phases. Firstly, it identifies the blockage events when the variations of DL SNR surpass a threshold. Then, it recovers the blocked beams after establishing an alternative beam pair with a neighbour BS, and finally, it switches back to the primary beam when the DL SNR is back to the values without blockage. We assume that the quality of the beam l, forming the primary beam pair, can be tracked using both the instantaneous SNR, expressed as $\mathrm{SNR}_l(t) = \frac{P_b \left| \mathbf{b}_{rx}^{\mathsf{T}} \mathbf{H}(t) \mathbf{b}_{tx} \right|^2}{\sigma_z^2}$ and the average value of SNR for the non-blocked time instants

denoted as $\overline{\text{SNR}}_l$. For example, in the NR Standard, the beam quality is obtained at the UE through the RSRP measurements, based on the Channel State Information - Reference Signals (CSI-RSs) reception. This is reported periodically back to the BS using the Uplink (UL) control channels [94, 131]. The RSRP values can be mapped to the SNR through a linear relationship. For instance, SNR[dB] = RSRP[dBm] + 122 dBm, when the receiver noise is −122 dBm.[1]

In mmWave indoor networks, multiple BSs whose coverage areas overlap can provide alternative links if the primary link becomes blocked. Hence, the UE can connect to a neighbour BS, referred to as secondary BS, which provides the second-largest RSRP to the UE. The secondary BS may have a beam pair, referred to as backup beam pair, non-overlapping with the primary beam pair and not affected by the blockage when the primary is blocked, as depicted in Fig. 4.7a.

After the blockage detection, the UE access to the secondary BS and repeats the exhaustive search to identify the best beam pair [54]. Furthermore, we assume to repeat the beam alignment at each subsequent blockage event, since in the indoor environment, due to rotation of the device, even the pre-computed backup beams may become blocked, which requires frequent re-initialisation of the alignment [81]. Additionally, as shown in Fig. 4.6, the SSBs use limited spectrum resources, and the remaining spectrum resources may be dedicated to data transmission [54]. Thus, we assume that data transmission occurs in the TTIs used for the SSB transmissions.[2]

Fig. 4.7b represents an example of the primary and backup beams for one of the network's UE and their respective SNRs. We define as Th_1 and Th_2 two thresholds with $Th_1 < Th_2$. We fix Th_1 and Th_2 values corresponding to $\overline{\text{SNR}}_l$ minus 70% and 30% of the maximum blockage loss BL_{max}, which is in line with the procedure adopted in [89]. Whenever the beam is non-blocked, the SNR shows small fluctuations around the average value $\overline{\text{SNR}}_l$ due to the channel's multipath components. Let us denote by $\bar{t}$ the time-instant when the

[1] The mapping between RSRP and SNR is valid under the assumption that the RSRP – computed as the average power of all the cell-specific CSI-RSs carried over multiple RBs – does not account for the inter-cell interference.

[2] The data transmission in the SSB TTIs may be possible under certain circumstances. For instance, the secondary BSs may serve those UEs in the same direction of the SSB transmission [54] whereas the serving BSs may schedules the UE in different TTIs from the ones used by the SSBs or the UE may perform beam sweeping within the TTI used for SSB transmissions.

blocker starts to intersect the LoS path between serving BS and UE; at this point, the SNR starts to decay rapidly, and the blockage event is detected at any time instant t if $\mathrm{SNR}_l(t) < Th_1$. A procedure follows this event to align the backup beam pair. The time interval between the start of the blockage and the time instant when the BR-Det method switches to the backup beam pair can be expressed as

$$\beta_1 = t_{d1} + T_{sweep} + T_{HO}, \tag{4.11}$$

where $t_{d1} > 0$ is the interval between $\bar{t}$ and the time instant when the SNR decays below the threshold Th_1, T_{sweep} denotes the exhaustive search duration to align the backup beam pair and T_{HO} represents the time interval before handover to the secondary BS.

Even the secondary BS uses beam sweeping to serve one UE at a time, including the UE that switches to the backup beam pair. The data rate of the DL transmission between the secondary BS j' and UE k can be expressed as

$$r''_{j',k}(t) = \frac{BW}{K_{j'}} \log_2 \left(1 + \frac{P_b \left| \bar{\mathbf{b}}_{rx}^{\mathsf{T}} \mathbf{H}_{j,k}(t) \bar{\mathbf{b}}_{tx} \right|^2}{I_{j',k}(t) + \sigma_z^2} \right), \tag{4.12}$$

where $\bar{\mathbf{b}}_{tx}$ and $\bar{\mathbf{b}}_{rx}$ denote the TX and RX beamforming vectors forming the backup beam pair with BS j'.

At the end of the blockage event, which lasts for the duration T_{bl}, the SNR sharply increases to values around $\overline{\mathrm{SNR}}_l$ and, at any time instant, if $\mathrm{SNR}_l(t) > Th_2$, we declare the beam has returned to the non-blocked state, and a procedure follows to align with the primary beam. The time interval between the end of the blockage and the time instant when the BR-Det method switches back to the primary beam pair can be expressed as

$$\beta_2 = t_{d2} + T_{sweep} + T_{HO}, \tag{4.13}$$

where $t_{d2} < 0$ is the interval between the time instant when the blockage ends, i.e. $\bar{t} + T_{bl}$, and the time instant when the $\mathrm{SNR}(t) > Th_2$, T_{sweep} and T_{HO} are assumed to be the same as in Eq. (4.11). The data rate of UE k obtained while adopting the BR-Det method can

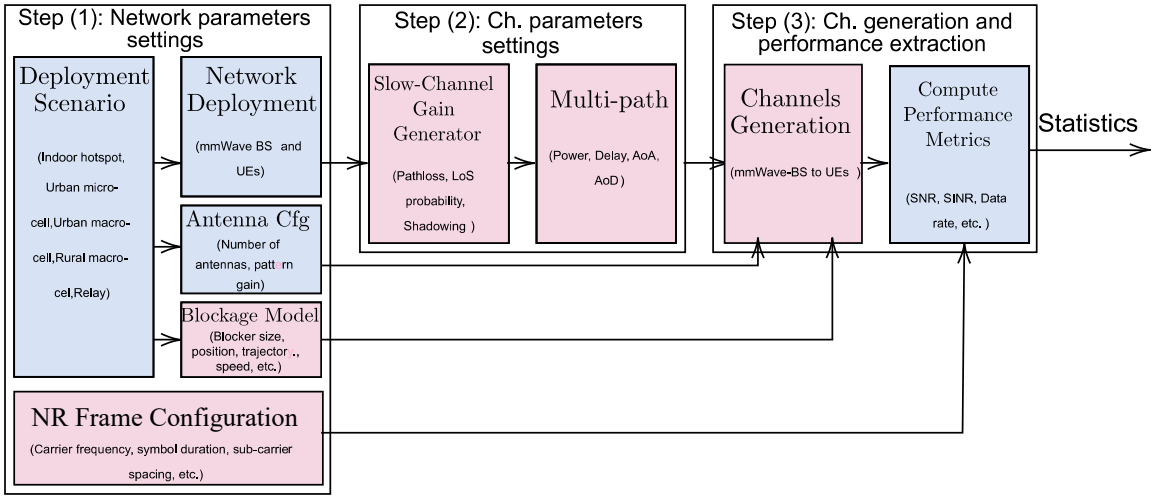


Figure 4.8: Illustration of the system-level simulator components, including the ones highlighted with the red background colour that simulate 3D SCM described in Sec. 4.2.2 and the Blockage Model B described in Sec. 4.2.3.

be expressed as

$$r_k^{\mathrm{BRDet}}(t) = \begin{cases} r'_{j,k}(t) & \text{for } 0 < t < \bar{t} + \beta_1, \\ r''_{j',k}(t) & \text{for } \bar{t} + \beta_1 \leq t < \bar{t} + T_{bl} + \beta_2, \\ r'_{j,k}(t) & \text{for } t \geq \bar{t} + T_{bl} + \beta_2. \end{cases} \tag{4.14}$$

4.3 Simulations and Numerical Results

In this section, we first describe the setup of the 3GPP-based system-level simulations. Then, we present the time response of the BR-Det method. Finally, we show the data rate and latency performance of the mmWave network when adopting the BR-Det method.

4.3.1 Simulation Setup

We simulate the network performance considering a drop-based system-level simulator illustrated in Fig. 4.8. This version of the system-level simulator is a major upgrade of the one described in Sec. 3.5.1, as it integrates the most recent 3GPP 3D SCM for modelling the mmWave channel, the geometric blockage Model-B and the NR frame numerology [57, 94], which are all represented in Fig. 4.8 with blocks having the red background colour. The simulator is calibrated accordingly to the 3GPP procedure described in [57, Sec. 7.8] that

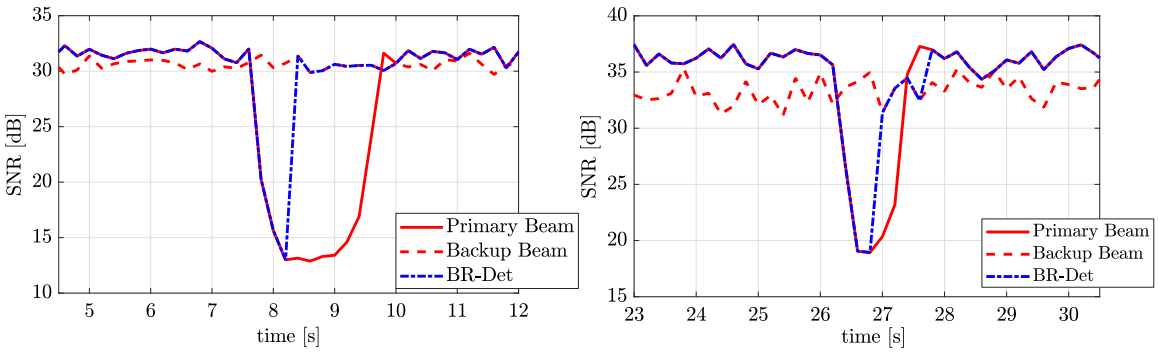
Table 4.1: 3GPP-based system-level simulation parameters

NR Numerology	**Description**
Carrier frequency	28 GHz
Nominal system bandwidth/Number of RB	396 MHz/275 RB [94]
Sub-carrier spacing/Slot-duration	120 KHz/0.125 ms [94]
BS	**Description**
Network layout	Room size of 40 m × 50 m × 3 m, 4 sites, 3 sectors/site
Deployment	Grid-based with ISD: 20 m, height: 3 m
Antenna array	UPA with element spacing 0.5λ, Number of antennas per array: $8 \times 8 = 64$, Mechanical downtilt: 20°, Antenna boresigh $\{30°, 150°, 270°\}$ [126]
Single antenna element pattern	90° H × 90° V beamwidths, 5 dBi max. [126]
Beamforming	Fully-analog architecture based on 2D codebook $N_{\mathrm{CB},tx} = 64$
TX power	20 dBm
UE	**Description**
Deployment	Random, 20 UE/sector on average, all UE served, height: 1 m
Antenna array	UPA with element spacing 0.5λ, Number of antennas per array: $4 \times 4 = 16$, Antenna boresigh $\sim \mathcal{U}[0°, 360°]$ [126]
Single antenna element pattern	90° H × 90° V beamwidths, 5 dBi max. [126]
Beamforming	Fully-analog architecture based on 2D codebook $N_{\mathrm{CB},rx} = 16$
Noise figure	10 dB
Blocker	**Description**
Model	Blockage Model B [57]
Dimensions	2 m × 3 m
Trajectory/Speed	Linear, $v = \{1, 2\}$ m/s
Channel	**Description**
Path loss and LOS probability	3GPP 3D InH-Open office [57]
Shadowing	Log-normal with $\sigma = \{3, 8\}$ dB (LoS/NLoS) [57]
Fast fading	Ricean with log-normal K-factor [57]
Thermal noise	−174 dBm/Hz power spectral density

uses the list of parameters indicated in [57, Table 7.8-2].

The network deployment is formed by BSs placed on the rectangular grid and UEs distributed uniformly in the scenario. We consider one blocker at a time moving in the scenario, with dimensions 2 m × 3 m and speed selected between two values 1 m/s and 2 m/s, in line with the recommended blocker parameters described in [57]. The initial blocker position is set at the x-y coordinates $(-20, 0)$ m, corresponding to the left side of the network layout, and it moves along the x-axis line from left to right of the network layout. When the blocker reaches the right side of the scenario, it regenerates to the starting position and repeats the same movement. We configure the simulator accordingly to the list of parameters of Table 4.1.

We simulate 10 network drops in total, where each network drop corresponds to an in-



(a) SNR time series for blocker speed 1 m/s.

(b) SNR time series for blocker speed 2 m/s.

Figure 4.9: Evolution in time of the BR-Det method SNR for two different beams and blocker speeds.

dependent deployment of BSs, UEs and blockers. The UEs locations are assumed not to change between drops whereas their antenna orientation changes. Between network drops, we consider time-invariant large-scale channel parameters (LSPs), such as delay spread, Direction of Arrival (DoA) and AoA spread, Rician K-factor, shadow fading (SF) and pathloss, for the duration of the drop T. Additionally, at each simulation step within a drop, we assume *i)* time-varying blockage loss conditions due to the blocker's movement and *ii)* time-varying realisations of small-scale channel parameters (SSPs), e.g. power, phase, delay, AoA and AoD for each individual sub-path and cluster, due to the variations of the multipath fading. We fix the network drop duration to $T = 40$ s, and we consider an interval $\Delta t = 200$ ms to update the channel $\mathbf{H}$ between the simulation steps.

Accordingly to the TX and RX codebook dimensions, that are $N_{\mathrm{CB},tx} = 64$ and $N_{\mathrm{CB},rx} = 16$, we account for the time required to complete the exhaustive search that is $T_{sweep} = 330$ ms. Additionally, we fix the time to perform the handover at $T_{HO} = 50$ ms [132].

4.3.2 Time Response Analysis of the BR-Det Method

Fig. 4.9 shows the evolution of the SNR of the primary beam, the backup beam, and their combinations when adopting the BR-Det method with the blocker moving at the speeds 1 m/s and 2 m/s.

Looking to the Fig. 4.9a, the SNR of the primary beam starts decreasing due to the blocker crossing the beam path. Simultaneously, the backup beam shows a lower SNR than

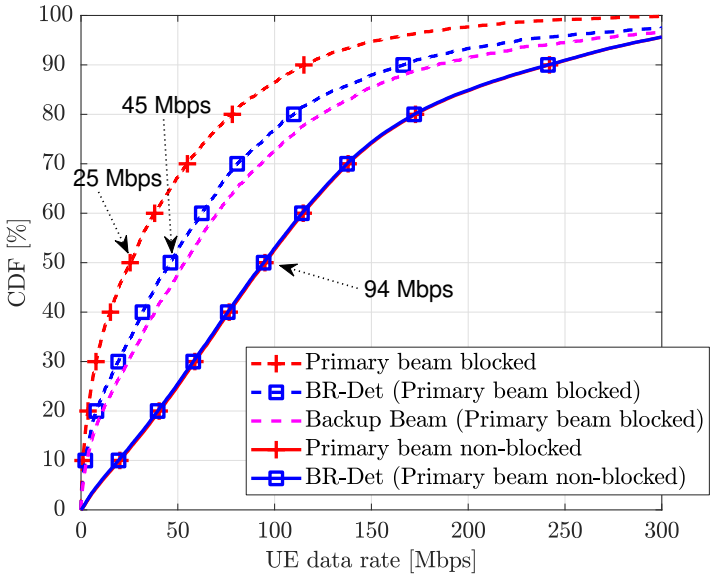
the primary beam because it is established with a secondary BS. However, the SNR remains steady around its average value for the entire blockage duration since it uses the backup beam that is not affected by the blockage when it crosses the primary beam. The BR-Det method switches to the backup beam when *i)* the SNR decays below the threshold Th_1 and *ii)* the beam sweeping to establish the backup beam ends.

Fig 4.9b represents similar behaviour. Nevertheless, while comparing Fig. 4.9a and Fig. 4.9b, we remark that the BR-Det method's delay to switch from the primary to the backup beams is the same in the two figures and corresponds to $\beta_1 = 600$ ms. On the other hand, the blockage duration changes between 2 s and 1 s. As a result, the ratio between the BR-Det method response time and the blockage event duration increases from 30% to 60%, which indicates that the BR-Det method takes 30% and 60% of the blockage duration to detect the blockage and switch beam.

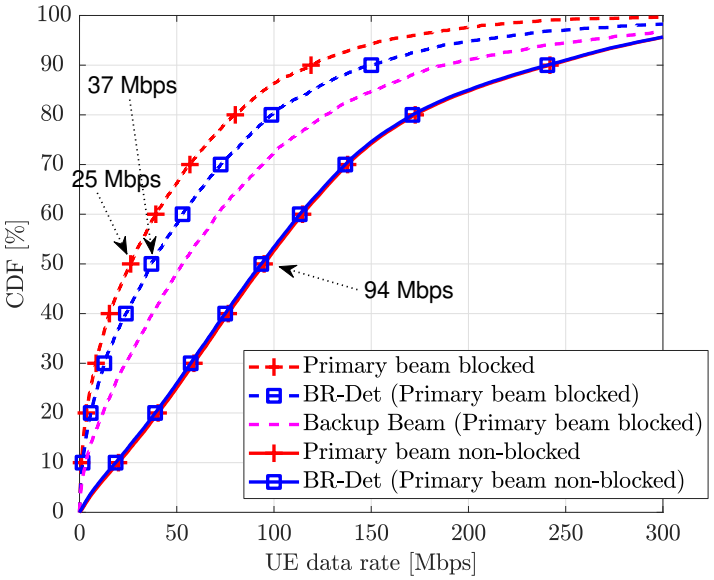
4.3.3 Data Rate Performance of the BR-Det Method

In this section, we show the Cumulative Distribution Function (CDF) of the UEs data rate for the BR-Det method with the blocker moving at speeds 1 m/s and 2 m/s. We maintain the same blockage conditions, i.e. one blocker at a time and the same simulation duration, for the different blocker speeds. Thus, the blockage probability P_b, which considers the total duration of the blocked intervals over the total simulation time, remains constants. We compare the BR-Det method results in the other two cases: *i)* when UEs use only the primary beam, and there is no beam recovery, and *ii)* when UEs switch to the backup beam pair without delay, i.e. $\beta_1 = 0$. We further differentiate the results between the blocked and non-blocked time intervals.

Fig. 4.10a shows the CDF of the UEs data rate for the blocker speed of 1 m/s. The median data rate decreases from 94 Mbps to 25 Mbps. During the non-blocked intervals, the signal-to-interference-and-noise ratio (SINR) is high, and the UE data rate achieves good performance. In contrast, during the blocked intervals, the SINR decreases considerably for the blockage effect, and the UEs data rate is severely penalised. At the same time, the BR-Det method results show a median data rate of 45 Mbps, which improves by 80% the median data rate of the primary beam. This result is also close to the backup beam's



(a) CDF of UEs data rate for the blocker speed 1 m/s.



(b) CDF of UEs data rate for the blocker speed 2 m/s.

Figure 4.10: UE data rate performance for blocker speeds 1 m/s and 2 m/s.

median data rate, where we assume switching to the backup beam pair without delay.

Secondly, in Fig. 4.10b, we show the CDF of the UEs data rate for the blocker speed of 2 m/s. Here the primary and backup results match the ones shown with blocker speed 1 m/s because the blockage conditions are the same in the two figures. Nevertheless, the BR-Det method results show a median data rate of 37 Mbps, which improves by 48% the primary beam's median data rate. To note that the gain of using the BR-Det method reduces with the higher blocker speed. This is explained by the fact that the higher blocker

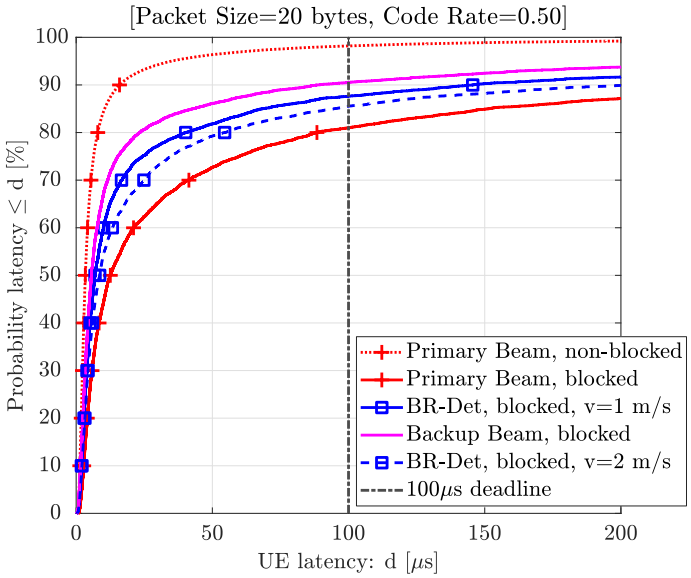
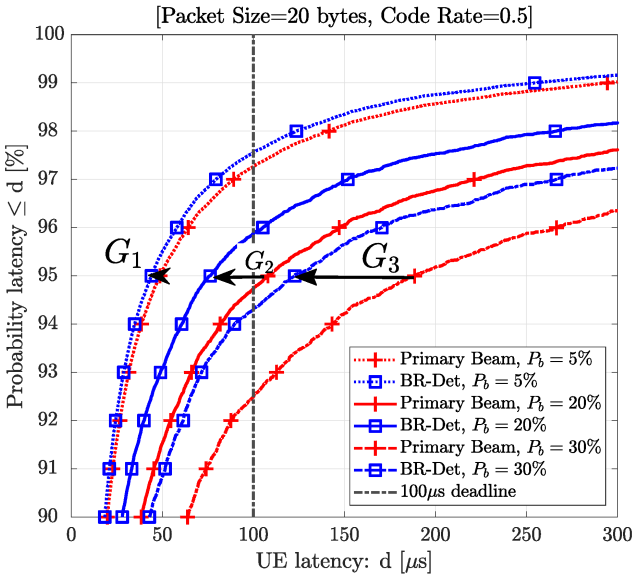


Figure 4.11: CDF of UEs latency for Primary, Backup beams and BR-Det method.

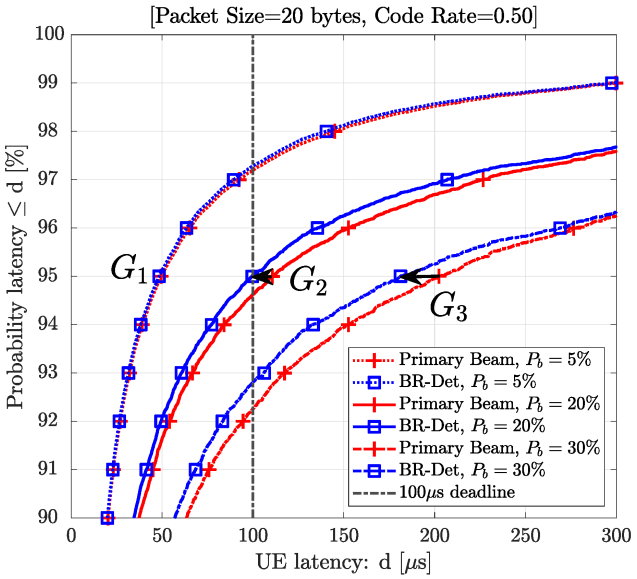
speed increases the frequency of blockage events since the blocker moving indoor travels for a longer distance and may intersect more times the same beam. For every blockage event, the BR-Det method switches to the backup beam with a delay, as shown in Sec. 4.3.2, penalising the BR-Det data rate at the beginning of the blockage event. Therefore, as the BR-Det method switches to the backup beam more frequently, the BR-Det method's data rate is more penalised with a blocker speed of 2 m/s than with a blocker speed of 1 m/s.

4.3.4 Air Interface Latency Performance of the BR-Det Method

In this section, we present the air interface latency results, which measure the delay for transmitting a fixed-size packet over the air interface. Let us suppose that the real-time control of a moving robot requires 1 ms end-to-end latency to transmit a packet of P_{inf} bytes, which includes the overhead due to the channel coding. Such end-to-end latency requires air interface latency not larger than 100 μs per direction [69], leaving the rest of the time budget for other operations [69]. In this example, we consider transmitting a packet with a size of 20 bytes [70], and we account for the channel coding overhead given by a code rate of 1/2 [133]. Thus, the air interface latency can be expressed as $d_k(t) = \frac{P_{enc}}{r_k(t)}$, where $P_{enc} = 8 \times P_{inf} C_r$ is the encoded packet length in bits, C_r is the code rate and $r_k(t)$ is the data rate.



(a) 95-th percentile latency for blocker speed 1 m/s.



(b) 95-th percentile latency for blocker speed 2 m/s.

Figure 4.12: 95-th percentile latency varying the blockage probability.

Fig. 4.11 presents the CDF of the UEs latency values. The percentage values of the CDF reported on the y-axis indicate the probability that the latency is lower than a deadline shown on the x-axis. The higher the data rate is, and the higher is the probability that the packet is transmitted in a given time interval. As shown in Fig. 4.11, the median latency value (50-th percentile) is significantly lower than the deadline of $d_{\text{PHY}} = 100$ μs for both blocked and non-blocked intervals. Nevertheless, the other latency values largely spread around the median value. A typical spread parameter, also considered in [70], is the

value of the 95-th percentile, i.e. the maximum latency value for the 95% of the latency values. Fig. 4.11 shows that the maximum value for the 95-th percentile of the non-blocked case complies with the deadline $d_{\mathrm{PHY}} = 100$ μs, meaning that the primary beam without blockers would operate with a latency lower than the PHY deadline for most of the cases. On the other hand, the primary beam's 95-th percentile latency during the blocked intervals is less than the deadline in 81% of the cases. There is a 19% probability that the latency may exceed the deadline during the blocked intervals. To better analyse this situation, we compare in Fig. 4.12, the 95-th percentile of the CDF varying the blockage probability P_b and considering both blocked and non-blocked time intervals. Fig 4.12a shows the 95-th percentile latency results with the blocker speed $v = 1$ m/s. The latency value increases rapidly, increasing the link's blockage probability and the primary beam latency results exceed the deadline with $P_b = 20\%$ and $P_b = 30\%$. The BR-Det method reduces the primary beam latency results by $G_2 = 30\%$ and $G_3 = 35\%$, when the blockage probabilities are $P_b = 20\%$ and $P_b = 30\%$, respectively, and the latency becomes lower than the PHY deadline when $P_b \leq 20\%$.

Conversely, with the blocker moving at speed $v = 2$ m/s as depicted in Fig 4.12b, the BR-Det method reduces the primary beam latency results by $G_2 = 10\%$ and $G_3 = 11\%$, when the blockage probabilities are $P_b = 20\%$ and $P_b = 30\%$, respectively, and the latency complies with the PHY deadline if $P_b \leq 20\%$. Overall, the gains with the blocker speed 2 m/s are lower than the gains with a blocker speed of 1 m/s. The BR-Det method performances suffer from a lower data rate, as shown before in Sec. 4.3.3, due to the initial delay happening more often as the BR-Det method switches to the backup beam more frequently with the higher blocker speed. Consequently, the air interface latency increases and the BR-Det method does not improve the system's performance without recovery as long as the blocker moves at the higher blocker speed.

4.4 Conclusion

In this chapter, we analysed the performance of a mmWave indoor network under dynamic blockage effect. We used a 3GPP-compliant geometric blockage model, which represents the blocker as a 3D rectangular screen and computes the total power loss assuming knife-edge

diffraction from the edges of the screen given the positions of blocker, RX and TX.

The mmWave link formed by the primary beam pair aligned between BSs and UEs shows rapid and large received signal power drops when it encounters a blocking object moving in the environment. We adopted the BR-Det method that detects when the blocker enters the primary beam path and searches for a backup beam pair, which provides a temporarily stable link while the primary beam is blocked. Thus, we tested the BR-Det method data rate performance in the mmWave indoor network scenario, varying the blocker speed.

This study's first takeaway is that the blocker significantly impacts the primary beam pair's data rate during the blocked intervals. Adopting the BR-Det method increases the median data rate of the primary beam by 80%. However, this improvement is considerably lower, varying the blocker speed to 2 m/s, where the BR-Det method improves by 48% the primary beam results. The delay of the beam recovery operations turns out to penalise more the BR-Det method's data rate with the higher blocker speed. This happens because the blocker moving indoor at a higher speed creates more blockage events that increase the beam switching frequency. Therefore, the data rate losses at the beginning of the blockage events occur more often.

As a second takeaway, we showed that UEs' latency value during non-blocked intervals is included within the PHY deadline for the 95-th percentile cases. Nevertheless, the presence of the blocker increases the 95-th percentile latency value considerably. Without deploying the beam recovery method, the 95-th likely latency exceeds the deadline when the blockage probability is larger than 20%. Under the same blockage probability conditions and by adopting the BR-Det method, the 95-th percentile latency becomes lower than the PHY deadline. Similarly to the data rate, the latency reduces more with the blocker speed of 1 m/s than the blocker speed of 2 m/s.

To summarise the above discussion, we provide the final remarks. The BR-Det method improves the system data rate and latency performance for the blocker speed of 1 m/s. However, it shows some limitations for the blocker speed of 2 m/s, where the frequency of the beam switching operations severely penalises the BR-Det method.

This behaviour represents a potential problem for applying the BR-Det method to the mmWave indoor networks because it can be adversely affected by the blockers moving faster.

Consequently, there is a need to investigate alternative methods to the detection-based method to make beam recovery operations more effective. The next chapter of this dissertation will look deeper into this aspect.

5 Predictive Indoor mmWave Networks with Dynamic Blockers

In this chapter, we discuss a millimetre Wave (mmWave) network that uses Machine Learning (ML) models predictions into the network control operations to assist the network in taking the beam switching decisions. The technical work presented in this chapter is based on our paper [21].

5.1 Introduction

The integration of ML tools into wireless communication systems is envisioned as one of the essential steps towards the realisations of future intelligent Radio Access Networks (RANs) [13, 134]. At the same time, ML tools can also optimise the resource management and control operations of the 5-th Generation (5G) networks to render these systems extremely efficient and enhance their performance [14]. One of the key technologies for 5G is mmWave communication, which enables transmission data rates in the order of Gb/s to accommodate new services, e.g. Enhanced Mobile BroadBand (eMBB) and Ultra-Reliable Low-Latency Communication (URLLC) [4, 10, 121]. Nevertheless, the high penetration loss and low diffraction of mmWave frequencies make this technology very sensitive to objects passing through the mmWave links, causing temporal blockages of the signal [122]. When the Transmitter (TX)-Receiver (RX) beam pair becomes obstructed, e.g. by moving blockers, the received signal power experiences a significant attenuation. In addition, the link's data rate may become highly intermittent due to the dynamic blockers moving in the environment, making it extremely challenging to provide high data rate and low latency connections during the blockage events [51, 78].

The conventional beam recovery methods based on detection, e.g. the one presented in Sec. 4.2.6 and other techniques introduced in [81, 82], rely on a signal threshold that detects the blockage events only when the received signal power drops below a given value. The beam recovery operations yield a delay as they establish the backup beam after detecting the blockage, leading to a data rate loss every time the blocker intersects the mmWave link.

A different approach consists of predicting the start and ends of blockage events to switch preemptively to a backup beam in case of blockage events [88]. The predictions can be provided by "sensing" the environment surrounding the mmWave link, for example, by using other mmWave links [11, 87], utilising the Sub-6 GHz spectrum bands [12, 135] or by adopting external sensors (e.g. camera [91, 92] or radar [93]) that needs to be interfaced with the mmWave network. Thus, leveraging these data makes it possible to predict if a blockage event will occur at any given time on the mmWave link. According to the blockage predictions, triggering the beam recovery operations ahead of the blockage events can avoid the data rate loss happening with the detection methods. These early works [11, 12, 87, 91, 92] showed promising results when applying ML tools for the blockage prediction problem. Nevertheless, their analysis is limited to the single-User Equipment (UE) scenario, and it does not account for multiple UEs. The multi-UE scenario represents a more realistic scenario for mmWave network deployments and enables exploring the feasibility to apply ML predictions at the network level.

5.1.1 Scope and Contributions

In this chapter, we propose using Deep Neural Network (DNN) models' predictions into a novel beam recovery method that anticipates the beam switching operations in case of imminent blockage. The predictions indicate the start and finish of blockage events on a specific beam and are provided by a set of DNN models trained using data reported by other beams through the network protocol measurements. These data indicate whether large variations of the mmWave channels for those beams intersect the blocker in locations spatially close to the beam to predict. Since the beam-quality measurements follow similar temporal dynamics for beam spatially close, they may show a large drop of the signal-to-noise ratio (SNR) before it also happens on the beam to predict. Hence, the blocker presence

at time t on other beams can indicate the blockage presence on one beam in successive time instants. Thus, we train a DNN model to predict the beam-specif blockage state, taking as input the other beams' measurements exchanged with the neighbour Base Stations (BSs) through a central server. Unlike the works [11, 12, 87, 91, 92], we consider extending the blockage prediction methods to a more realistic multi-UE setting, which enabled us to deploy the ML models into the network and study its performance. Moreover, most of the existing works necessitate extra Hardware (HW) components, e.g. camera and Sub-6 GHz transceiver, that need to interface with the mmWave network operations. Instead, our approach uses the beam-quality measurements reported by the UEs to the BS that are already present in the Third Generation Partnership Project (3GPP) New Radio (NR) standard specification [94].

We experiment this approach by considering a mmWave Indoor network with high UEs density and multiple BSs. We account for BSs employing multi-antenna arrays with analog beamforming architecture and codebook-based signal transmission. Firstly, we generate synthetic data with a 3GPP-based system-level simulator incorporating a geometric blockage model [57]. We train beam-specific DNN models with this dataset to predict the blockage states of different beams of the codebook. Our contributions can be summarised as follows:

1) We propose a novel multi-UE prediction-based method for beam recovery that incorporates the beam measurements from multiple BS to predict the subsequent blockage events. This early indication enables to initiate the beam switching in advance and complete this operation by the time the mmWave link becomes blocked.
2) We validate the prediction-based method by deploying the DNN models online into a 3GPP-based system-level simulator to verify the close match to the performance of an ideal method that has perfect knowledge of the future beam states. We then compare the prediction method to other two cases: *i)* adopting the beam recovery method based on detection and *ii)* utilising a fixed beam method. Finally, we show that the prediction method switches to a backup beam earlier than the blockage event, and avoids the data rate loss that occurs with the detection method due to the delay in switching to the backup beam and without the switching for the fixed beam method.

3) We provide a quantitative analysis of the three methods' data rate performance with different blocker speeds.

The remainder of this chapter is organised as follows. Section 5.2 describes the system model and presents the beam recovery based on the prediction method; Section 5.3 presents the procedure used to obtain the beam state predictions; Section 5.4 presents the main results of the analysis, and Section 5.5 summarises the key findings.

Notation: The following notation is used throughout the chapter: boldface lower case and boldface upper case are used for column vectors $\mathbf{x}$ and for matrices $\mathbf{X}$, respectively. $\mathbf{x}^{\mathsf{T}}$ denote the transpose of $\mathbf{x}$. A Gaussian random variable x is denoted $x \sim \mathcal{CN}(\mu, \sigma^2)$, where μ is the mean and σ^2 is the variance. $[X]$ denotes the Iverson bracket defined as 1 when X is true, and 0 when X is false.

5.2 System Model

We consider the same mmWave network introduced in Sec. 4.2.1, which is formed by four cell sites deployed on a rectangular grid with an Inter Site Distance (ISD) of 20 meters covering an area of 50 meters by 40 meters, as depicted in Fig. 4.1. We recall that a mobile blocker with dimension $w \times h$ moves at a constant speed v following a linear trajectory at the centre of the environment. Each cell site has three sectors, served respectively by three BSs oriented with an angle θ_j with $j \in \{1, 2, 3\}$ and operating at mmWave carrier frequency f_c, with bandwidth BW. All the BSs in the mmWave network form the set $\mathcal{J}$ with cardinality J. Additionally, we consider a set of $\mathcal{K}$ UEs distributed uniformly at random in the mmWave network area.

We assume that each BS (TX) and UE (RX) employ Uniform Planar Array (UPA) with antenna spacing of a half-wavelength and number of antennas $M_{tx} = M_{tx}^{\mathrm{V}} \times M_{tx}^{\mathrm{H}}$ and $M_{rx} = M_{rx}^{\mathrm{V}} \times M_{rx}^{\mathrm{H}}$, where the subscripts V and H represent the vertical and horizontal dimensions of the UPAs. Both the BS and the UE use analog beamforming, with a single Radio Frequency (RF) chain that enables to form narrow beams in different spatial directions accordingly to beamforming vectors $\mathbf{b}_{tx}$ and $\mathbf{b}_{rx}$, which are expressed in Eq. (4.6) and Eq. (4.7), and are taken from the codebooks $\mathcal{B}_{tx}$ and $\mathcal{B}_{rx}$ having sizes $N_{\mathrm{CB},tx}$ and $N_{\mathrm{CB},rx}$.

5.2.1 Downlink Data Transmission

Consider the Downlink (DL) of a NR-based system for mmWave cellular communications operating with the Time Division Duplex (TDD) scheme. In the initial access, the UE performs a beam sweeping procedure with one of the neighbouring BSs through an exhaustive search to identify the best beam pair, referred to as the primary beam pair, which provides the maximum Reference Signal Received Power (RSRP). The overhead time to complete the exhaustive search is defined as in Sec. 4.2.5 and is computed as $T_{sweep} = T_{SS}\frac{N_{\mathrm{CB},tx}N_{\mathrm{CB},rx}}{L_{\mathrm{SSB}}} + T_{SS}/2$, where T_{SS} is the time interval to transmit one Synchronisation Signal (SS) Burst and L_{SSB} is the maximum number of synchronisation signal block (SSB) within each SS Burst.

During the data transmission, each BS uses beam sweeping to time-multiplex multiple UEs. The data rate of the DL transmission between BS j and UE k can be expressed as

$$r^{'}_{j,k}(t) = \frac{BW}{K_j}\log_2\left(1 + \frac{P_b\left|\mathbf{b}^{\mathsf{T}}_{rx}\mathbf{H}_{j,k}(t)\mathbf{b}_{tx}\right|^2}{I_{j,k}(t) + \sigma_z^2}\right), \tag{5.1}$$

where K_j represents the set of UEs served by the BS j, P_b is the transmit power of the BS, $\mathbf{H}_{j,k}(t)$ represents the 3D channel (see Sec. 4.2.2), which includes the blockage effect $BL(t)$ (see Sec. 4.2.3), $I_{j,k}(t) = \sum_{j'\in\mathcal{J}\setminus j}\sqrt{P_b}s_{k'}(t)\mathbf{b}^{\mathsf{T}}_{rx}\mathbf{H}_{j,k}(t)\widehat{\mathbf{b}}_{tx}$ represents the inter-cell interference from each interfering BS $j' \in \mathcal{J}\setminus j$, $\widehat{\mathbf{b}}_{tx}$ represents the TX beam directed from BS j' to UE $k' \neq k$ and σ_z^2 is the noise power at the UE receiver.

We consider each UE periodically feedback to the BS the beam-specific RSRP measurement [131], which reports the primary beam pair's quality. The RSRP is computed as the average power of the cell-specific Channel State Information - Reference Signals (CSI-RSs), which does not include inter-cell interference. Hence, there is a linear relationship between RSRP and SNR values expressed in the logarithmic scale. For instance, with thermal noise -174 dBm/Hz, $\Delta_f = 15$ kHz and RX noise figure 10 dB, the mapping is SNR[dB] = RSRP[dBm] + 122 dBm. To note that the SNR, expressed in the linear scale as $\mathrm{SNR}_l(t) = \frac{P_b\left|\mathbf{b}^{\mathsf{T}}_{rx}\mathbf{H}(t)\mathbf{b}_{tx}\right|^2}{\sigma_z^2}$, includes the beamforming vectors $\mathbf{b}_{tx}$ and $\mathbf{b}_{rx}$ forming the primary beam pair and having beam IDs l and q defined as $l = \{1,\ldots,N_{\mathrm{CB},tx}\}$ and $q = \{1,\ldots,N_{\mathrm{CB},rx}\}$, respectively.

5.2.2 Beam Recovery Based on Predictions

In this section, we present a new beam recovery method based on the prediction (BR-Pre) that switches to an alternative beam pair before the blocker affects the quality of the primary beam pair. First, let us define with $S_l(t)$ the ground truth state of the beam l at time t that can be expressed as

$$S_l(t) = \begin{cases} 0 \quad \text{(non-blocked)} & \text{for } 0 < t < \bar{t} \cup t \geq \bar{t} + T_{bl}, \\ 1 \quad \text{(blocked)} & \text{for } \bar{t} \leq t < \bar{t} + T_{bl}, \end{cases} \tag{5.2}$$

where $\bar{t}$ represents the time-instant when the blocker starts to intersect the path of the primary beam pair and T_{bl} represents the blockage duration.

At the time $t-\eta$, where $\eta > 0$ represents the prediction window, the BR-Pre method evaluates the beam state predictions $\widehat{S}_l$, indicating whether the blockage obstructs the primary beam path at time t. If the beam is predicted as non-blocked, i.e. $\widehat{S}_l(t-\eta) = 0$, the BR-Pre method continues using the primary beam pair. Differently, when the beam is predicted as blocked, i.e. $\widehat{S}_l(t-\eta) = 1$, the BR-Pre method uses this early indication to determine an alternative beam for switching the communication. Similarly to the beam recovery method based on the detection (BR-Det) described in Sec. 4.2.6, we consider that the backup beam pair can be identified via an exhaustive search with a secondary BS. We also consider that data transmission occurs in the Transmission Time Intervals (TTIs) used for the SSB transmissions. This is possible without frequency repetition as the SSBs use limited spectrum resources, and the remaining spectrum resources may be dedicated for data transmission [54].

At the end of the beam search, the BR-Pre switches to a backup beam pair and the data rate of the DL transmission between the secondary BS j' and UE k can be expressed as

$$r''_{j',k}(t) = \frac{BW}{K_{j'}} \log_2 \left(1 + \frac{P_b \left| \bar{\mathbf{b}}_{rx}^{\mathsf{T}} \mathbf{H}_{j,k}(t) \bar{\mathbf{b}}_{tx} \right|^2}{I_{j',k}(t) + \sigma_z^2} \right), \tag{5.3}$$

where $\bar{\mathbf{b}}_{tx}$ and $\bar{\mathbf{b}}_{rx}$ denote the TX and RX beamforming vectors forming the backup beam pair. When the beam state prediction is correct, i.e. $\widehat{S}_l(t-\eta) = S_l(t)$, the switching happens

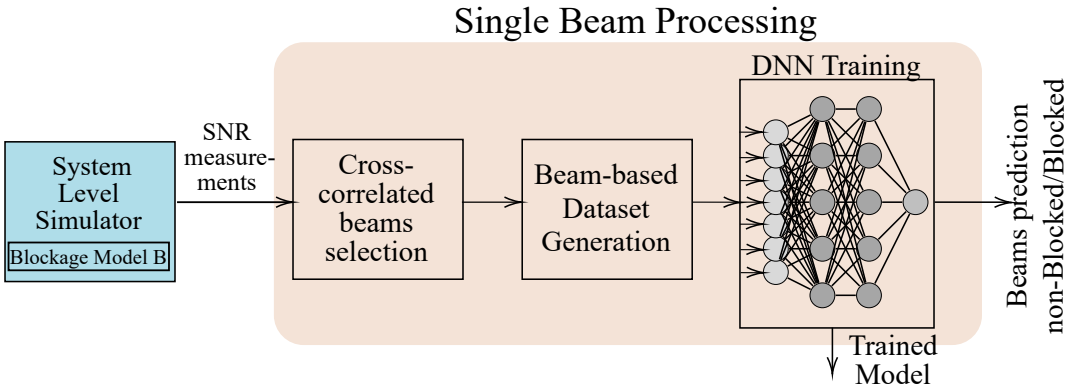


Figure 5.1: Key steps of the beam state predictions procedure. After generating synthetic data with the system-level simulator, the single beam processing consists of three steps: *i)* Selection of a subset of beams cross-correlated to the beam l, *ii)* Generation of a dataset formed by the SNR values of other beams (input) and the state of the beam to predict (output) and *iii)* Training of the beam-specific DNN model used to make the prediction.

before the blockage. While using the backup beam pair, the BR-Pre method continues to evaluate the predictions and if the beam l is predicted as non-blocked, i.e. $\widehat{S}_l(t-\eta)=0$, the BR-Pre switches back to the primary beam pair in t.

The data rate of UE k obtained while adopting the BR-Pre method can be expressed as

$$r_k^{\mathrm{BRPre}}(t)=\begin{cases} r'_{j,k}(t) & \text{if } \widehat{S}_l(t-\eta)=0, \\ r''_{j',k}(t) & \text{if } \widehat{S}_l(t-\eta)=1. \end{cases} \tag{5.4}$$

5.3 Beam State Predictions

In this section, we present a procedure used to obtain the beam state predictions $\widehat{S}_l$ for the BR-Pre method. The main idea is to use the SNR measurements of other TX beams – exchanged between BSs through a central server – to predict η time instants in advance the blockage state of the beam l at time t. As the blockage effect follows in general non-periodic temporal dynamics, the SNR does not change until the blocker intersects the beam l, making it unpractical to infer the beam-state only by analysing the SNR temporal variations of the single beam. On the other hand, by looking at the SNR temporal variations of other beams, their SNRs may drop earlier than the SNR of the beam to predict. We train a DNN model to relate these early SNR variations of other beams (inputs) to the blockage states of the beam l in successive instants of time (output).

Algorithm 1: Correlated beams selection for beam l

Input: A set $\mathcal{S} = \{\mathrm{SNR}_1(t), \ldots, \mathrm{SNR}_B(t)\}$ of time series defined in $0 < t < T$
Output: Set of cross-correlated beams

```
DelaysAll ← Vector(size: B − 1);
BeamsAll ← Vector(size: B − 1);
C_l ← Vector(size: L);
for l' ← 1 to B − 1 do
    for τ ← 0 to T do
        R_{l,l'}(τ) ← Compute cross-correlation;
    end
    δ = argmax_{∀τ} R_{l,l'}(τ) ; // Delay l-l'
    DelaysAll ← DelaysAll ∪ {δ};
    BeamsAll ← BeamsAll ∪ {l'};
end
Sort BeamsAll by the absolute value of DelayAll;
Copy the first L entries of BeamsAll into C_l;
return C_l;
```

Furthermore, as the BS is adopting a codebook with multiple beams, the DNN model needs to simultaneously provide multiple beam state predictions. For example, with a codebook formed by two beams, the first label indicates the state of the first beam (blocked, non-blocked) while the second label indicates the state of the second beam, assumed independent of the state of the first beam. Therefore, the task consists of learning one binary classifier per beam, meaning that we train and validate a beam-specific DNN model for each beam. As shown in Fig. 5.1, a sequence of steps leads to the beam-specific prediction $\widehat{S}_l$. We describe each step in the sections that follow.

5.3.1 Correlated Beam Selection

In this subsection, we introduce a pre-processing step that extracts a subset of beams to use for the predictions, considering the cross-correlation between the SNR time series of different beams. This procedure is motivated by two observations. First, the network presents many beams that will be irrelevant for predicting a specific beam state. This could needlessly complicate the training process without delivering any improvement. Second, since correlated beams are most likely spatially close, they also belong to the same BS or neighbour BSs. Therefore, we can also optimise the exchange of information among BSs to limit the use of a central server.

Let us consider the pair of TX beams l-l'. The delay between l and l' is estimated as $\delta_{l,l'} = \text{argmax}_{\forall\tau} R_{l,l'}(\tau)$, where $R_{l,l'}(\tau) = \sum_{t=0}^{T} \text{SNR}_l(t)\text{SNR}_{l'}(t-\tau)$ is the cross-correlation between the SNR time series of l and the SNR time series of l' defined over the time window $[0, T]$, where T is the total observation time. We estimate the delays between l and all the other beams $l' \in \mathcal{B} \setminus l$, where $\mathcal{B} = \{1, 2, \ldots, (N_{\text{CB},tx} \times J)\}$ represents the set of all the network's beams and $B = N_{\text{CB},tx} \times J$ is the total number of beams in the network. Then, we pick the L beams that have the lowest delays in absolute value, indicating that they are spatially close to beam l and we define a set $\mathcal{C}_l = \{c_1, c_2, \ldots, c_L\}$, which contains the beams to use for predicting the state of beam l. In Algorithm 1, we describe the procedure in details.

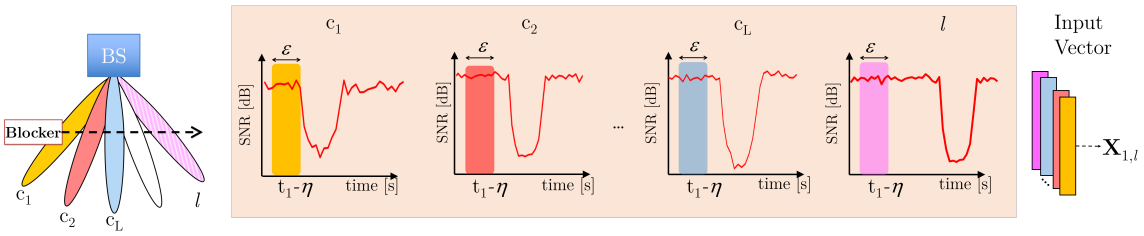
5.3.2 Dataset Structure

We consider a dataset formed by N samples, where the n-th sample of the dataset is organised into input/output vectors generated as follows:

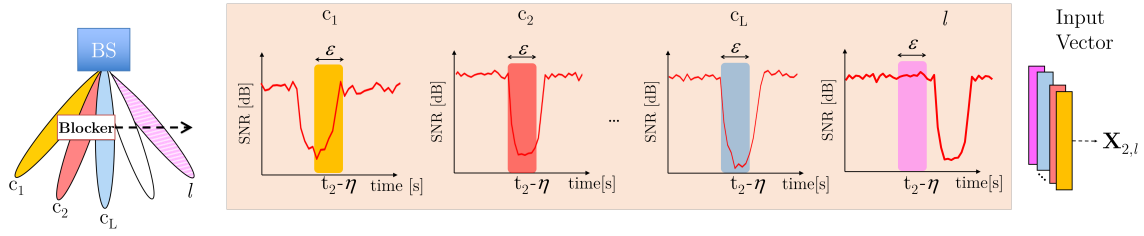
- **Input:** let us define as $\mathbf{v}_l = [\text{SNR}_l[t-\eta-(\epsilon-1)], \text{SNR}_l[t-\eta-(\epsilon-2)], \ldots, \text{SNR}_l[t-\eta]]^\mathsf{T}$ the vector containing ϵ samples of the SNR for beam l. We assume that the SNR measurements from the beams in the set $C_l = \{c_1, c_2, \ldots, c_L\}$ are exchanged with the BSs through a central server. These data are also represented for each beam $l' \in \mathcal{C}_l$ with a vector $\mathbf{v}_{l'} = [\text{SNR}_{l'}[t-\eta-(\epsilon-1)], \text{SNR}_{l'}[t-\eta-(\epsilon-2)], \ldots, \text{SNR}_{l'}[t-\eta]]^\mathsf{T}$ containing ϵ samples of the SNR. Thus, the input is defined as $\mathbf{X}_{n,l} = [\mathbf{v}_l, \mathbf{v}_{c_1}, \mathbf{v}_{c_2}, \ldots, \mathbf{v}_{c_L}] \in \mathbb{R}^{\epsilon \times (L+1)}$ and aggregates the SNR measurements of $L+1$ beams altogether.
- **Output:** the output is constituted by a variable $Y_{n,l}$ set to 1 (blocked) or 0 (non-blocked) according to the ground truth state of the beam l at time instant t. We obtain this information from the simulation environment considering the 3D rectangular screen modelling the blocker and the Line-of-Sight (LoS) path between BS and UE. If there is an intersection point, we set the label to 1 (blocked), otherwise 0 (non-blocked).[1]

During the beam-specific dataset generation, there are exceptions that we cover as follows.

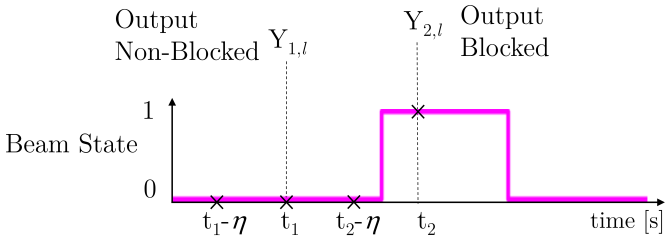
[1] We consider only the LoS path because from the simulations of this type of indoor scenario the probability of having a LoS link between the BS server and UE is 99.9%.



(a) Example of input data corresponding to the non-blocked ground-truth.



(b) Example of input data corresponding to the blocked ground-truth.



(c) Examples of output data for non-blocked and blocked beam states ground-truth for the magenta beam.

Figure 5.2: Two examples of input-output vectors forming the dataset for the beam depicted with the line pattern (magenta) on the left side of Figs 5.2a and 5.2b. In Fig. 5.2a, the input data are taken at a time $t_1 - \eta$ from the SNR measurements and correspond to the non-blocked state at the time t_1 in Fig. 5.2c . Conversely, in Fig. 5.2b, the input data are taken at a time $t_2 - \eta$ and correspond to the blocked state at the time t_2 in Fig. 5.2c

Firstly, the beam assigned to multiple UEs has an input vector filled with the median SNR and output variable corresponding to the most frequent label inside the UEs group. Secondly, if there are no UEs assigned to a beam, the corresponding SNR measurement cannot be recorded, and the input's entries are filled with a constant SNR (in our simulations, we use 60 dB) outside the range of the possible SNR values.

Figs. 5.2a and 5.2b show a simplified scenario where a set of beams, on the left side of the same BS, is used to predict the state of the beam l, depicted with the line pattern (magenta) on the right side of the BS. The entries of the dataset are formed by associating the inputs $\mathbf{X}_{1,l}$ and $\mathbf{X}_{2,l}$ to the values of the ground truth state $Y_{1,l}$ and $Y_{2,l}$ reported in Fig. 5.2c. The two different examples show how the input data changes depending on the value of the

Table 5.1: DNN network parameters

Network architecture	**Description**
Number of layers	2 (Fully-connected)
Number of hidden layer units	20
Weights initialisation	He
Activation function	ReLu
Dropout	None
Normalisation Input layer	Zero-center
Training options	**Description**
Solver name	Adam
Mini-batch size	1000
Number of Epochs	50
Learning rate	10^{-3}
L2 regularisation	10^{-5}

ground-truth. In the first example, depicted in Fig. 5.2a, the SNR measurements (within the sliding window of length ϵ) show small variations indicating that the set of beams at the time $t_1 - \eta$ is not affected by the blocker. This behaviour suggests that also the magenta beam is not likely affected by the blocker at the instant t_1. Indeed, as shown in Fig. 5.2c, the ground truth state corresponds to the non-blocked value, i.e. $Y_{1,l} = 0$. On the other hand, in the second example, depicted in Fig. 5.2b, the SNR measurements show large drops for the set of beams at the time $t_2 - \eta$, indicating that the magenta beam is likely affected by the blocker movement at the time t_2. This is confirmed by the ground truth state $Y_{2,l} = 1$, which corresponds to the blocked value.

5.3.3 Deep Neural Network Model Training

For training the ML model, we employ a DNN structure constituted by *i)* an input layer with $\epsilon \times (L + 1)$ inputs, *ii)* two hidden layers and *iii)* a `softmax` function as output layer. We train each DNN model with Adaptive moment estimation (Adam) optimiser adopting He weights initialisation and Rectified Linear unit (ReLu) activation functions [136, 137]. We also adopt a batch size of $N_{bs} = 1000$ samples and a constant learning rate for 50 epochs in total. Table 5.1 reports the training configurations and the list of hyper-parameters such as the number of hidden layer units, learning rate and L2 regularisation with the initial configuration values we adopted to start the training.

During the training, we divide the dataset into batches of size N_{bs} to find the set of model parameters denoted as $\mathbf{\Omega}_l$ that minimises the cross-entropy loss, given the set of inputs

$\mathbf{X}_{1,l}, \ldots, \mathbf{X}_{N_{bs},l}$ and the set of ground truth labels $Y_{1,l}, \ldots, Y_{N_{bs},l}$. The lower is the loss of the model output compared to the ground-truth and the better is the model. Moreover, since the time spent by the beam in the non-blocked state prevails over the time spent in the blocked state, the dataset presents a skewed distribution towards the non-blocked class that implies a class imbalance problem [138]. To re-balance the loss, we assign a higher weight to the cross-entropy loss computed for the rare class (blocked) and a lower weight to the cross-entropy loss calculated for the dominant class (non-blocked). The weighted formulation of the cross-entropy loss can be expressed as

$$\mathcal{L}_l(\mathbf{\Omega}) = -\frac{1}{N_{bs}} \sum_{n=1}^{N_{bs}} \mu_{1,l} Y_{n,l} \log(\widehat{Z}_{n,l}(\mathbf{\Omega})) + \mu_{2,l}(1 - Y_{n,l}) \log(1 - \widehat{Z}_{n,l}(\mathbf{\Omega})), \tag{5.5}$$

where $\widehat{Z}_{n,l}$ is the output of the `softmax` function that indicates the likelihood for the beam l to be in the state non-blocked or blocked. The weights $\mu_{1,l}$ and $\mu_{2,l}$ are computed as the inverse of the number of non-blocked and blocked training samples, i.e. $\mu_{1,l} = N/\sum_{n=1}^{N} [Y_{n,l} = 0]$ and $\mu_{2,l} = N/\sum_{n=1}^{N} [Y_{n,l} = 1]$, respectively.

5.4 Evaluation

In this section, we first describe the setup of the 3GPP-based system-level simulator used for generating synthetic data. We show the results of the DNN model state predictions for multiple beams on a training dataset. Then, we deploy the DNN models online into the simulator to study the BR-Pre method's performance. Finally, we compare the performance between the prediction and the detection methods.

5.4.1 Data Generation

We take a drop-based approach to generate synthetic data utilising a 3GPP-based system-level simulator described in Sec. 4.3.1. We assume the same simulation conditions adopted in Sec. 4.3.1 and the same set of simulation parameters indicated in Table 4.1. We recall that we consider one blocker at a time in the scenario, with dimensions 2 m $\times$ 3 m and speed selected randomly between two values 1 m/s and 2 m/s, in line with the recommended

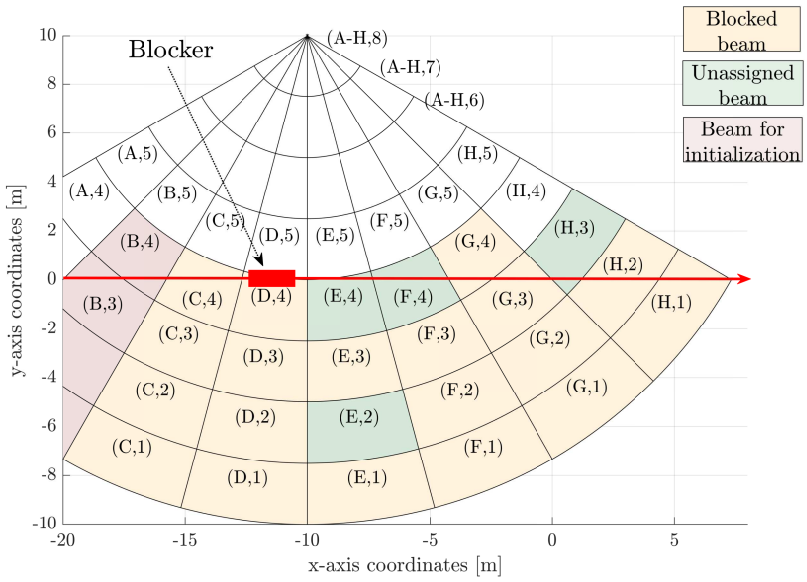
blocker parameters described in [57]. The initial blocker position is set at the x-y coordinates $(-20, 0)$ m, corresponding to the left side of the network layout, and it moves along the x-axis line from left to right of the network layout. We recall that when the blocker reaches the right side of the scenario, it regenerates to the starting position and repeats the same movement.

At each simulation step, we compute every BS-UE link received power, and we collect the DL SNR measurements for each beam, repeating the same process for 100 network drops. We configure the prediction window $\eta = 2$, which corresponds to 400 ms and covers both the exhaustive search duration that is $T_{sweep} = 330$ ms and the time interval before handover to the secondary BS that is $T_{HO} = 50$ ms [132]. We set $\epsilon = 10$, enabling us to track the SNR measurements through a sliding window of duration 2 s. Then, we use the number of input beams $L = 5$, which corresponds to a low DNN model complexity and reduces the amount of the measurement data exchanged with other BSs through a central server.

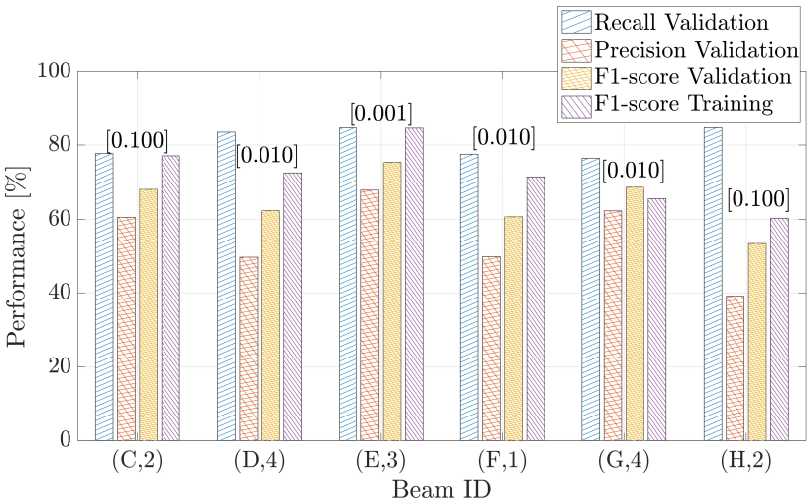
Subsequently, we generate a training dataset per beam formed by the collection of DL SNR measurements for all the 275 Resource Blocks (RBs) of the system bandwidth, for all the $189 = 200 - (\epsilon + \eta - 1)$ simulations steps and all the 100 network drops. We remove the samples without UE assigned to a beam from this large dataset. The remaining samples correspond to the blocked and non-blocked classes. For instance, the dataset of the beam $(\mathrm{C}, 2)$ has $\sim$1.5 million samples divided between $\sim$80 thousand samples for the blocked class and $\sim$1.42 million samples for the non-blocked class. We split this dataset according to the commonly used ratio of 80/20, which consists of using 80% of the data for training the DNN model and the remaining 20% of the data to validate the DNN model performance. Based on the validation results, we refine the model hyper-parameters, and we add the regularisation to each model independently. This reduces overfitting and improves the performance of the DNN model on the new data [139].

5.4.2 Multi-beam Prediction Results

In Fig. 5.3, we show the training and validation results of the DNN models for several beams of BS-3 corresponding to different azimuth and elevation Angle of Departures (AoDs) of the



(a) Illustration of the BS-3's beams crossed by the blocker.



(b) Validation results of the DNN models for blockage state prediction for several beams of BS-3.

Figure 5.3: DNN models results for multiple beams crossed by the blocker.

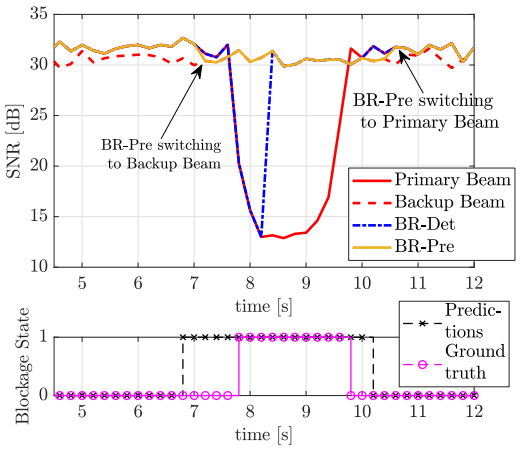
TX beamforming codebook. We use the BS-3 for its orientation, as it intersects the blocker multiple times with different beamforming directions. Nevertheless, the same process can be repeated without loss of generality for other BSs of the network.

Fig 5.3a shows the beams crossed by the blocker, the beams not assigned to UEs and the beam at the edge of the network layout used at the beginning of the simulation to initialise with the first samples of the sliding window. The DNN models of the blocked beams are validated considering the F1-score metric, which is more indicated than the Accuracy to measure the incorrectly classified cases when the class distribution is highly unbalanced

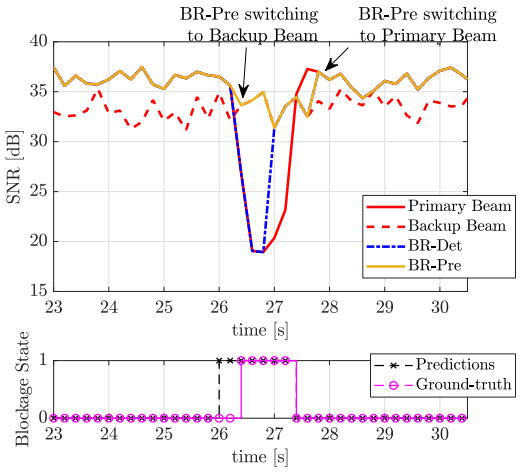
[138]. Likewise, in our scenario, the number of non-blocked class samples outweighs the blocked class ones. The F1-score combines through the harmonic mean Precision and Recall and can be calculated as F1-score $= \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$. Recall indicates the proportion of actual blocked samples (true positive (TP) cases) over all the blocked samples (TP and false negative (FN) cases), which are predicted correctly and is calculated as Recall $= \frac{\text{TP}}{\text{TP}+\text{FN}}$. Conversely, Precision indicates the proportion of actual blocked samples among all predicted blocked samples (TP and false positive (FP) cases) and is calculated as Precision $= \frac{\text{TP}}{\text{TP}+\text{FP}}$. In Appendix B, we provide a more detailed description of these metrics, and we report the confusion matrix, which helps visualise the different types of error.

The empirical analysis highlighted that the DNN models of different beams should be tuned separately. In other words, it is not possible to use the same set of hyper-parameters for all DNNs. Hence, we perform an independent validation of the DNN models beam by beam. During the validation, we change the L2 regularisation value in the range $\{10^{-5} - 10^{-1}\}$ to reduce the overfitting. We measure each beam's performance for five different random initialisation of the DNN model weights. Then, we save the model that returns the maximum value of the F1-score.

Looking at the results in Fig. 5.3b, we report between parenthesis the value of L2 regularisation hyper-parameter that corresponds to the model attaining the best F1-score performance on the validation data. Then, we report Recall, Prediction and F1-score validation results measured on the validation data and the F1-score training results measured on the training data. Firstly, all of the beams show Recall validation performance higher than Precision validation performance. This is due to the weighted loss function used during training, shown in Eq. (5.5), which penalises more the FN errors than the FP errors. This setting increases the probability of predicting the blockage correctly but is likely to make more FP errors for the non-blocked samples. Thus the models achieve high Recall while sacrificing their Precision. Secondly, Fig. 5.3b shows that the beam achieves F1-score validation results close to the F1-score training results. This indicates the models' ability to generalise well from the training data to any other data from the problem domain, as we will see in Sec. 5.4.4, where we deploy the DNN models online into the simulator.



(a) SNR time series and blockage state for blocker speed 1 m/s.



(b) SNR time series and blockage state for blocker speed 2 m/s.

Figure 5.4: Examples of observed SNR time series and blockage state predictions over a time window for two different beams and blocker speeds.

5.4.3 Evolution of the SNR Time Series

Fig. 5.4 shows the SNR measurements' evolution in time for two TX beams of the codebook with the blocker moving at the speeds 1 m/s and 2 m/s. These results highlight the BR-Pre behaviour over time and show the differences between the two beam recovery methods and their relationship with the blocker speed. Let us start to describe Fig. 5.4a. The bottom part of the figure shows the actual (ground-truth) and predicted primary beam states. It is worth recalling that the actual state is determined by considering the intersection between the direct path that joins TX and RX and the rectangular screen modelling the blocker, as described in Section 5.3.2. The upper part of the figure shows the corresponding evolution of the SNR of the primary beam pair, the backup beam pair, and their combinations resulting from the BR-Pre method's adoption or the BR-Det method. As shown in Fig 5.4a, the SNR of the primary beam pair starts decreasing when the beam is blocked, as depicted by the ground-truth state in the bottom part of Fig. 5.4a. Simultaneously, the backup beam pair shows a lower SNR than the primary beam pair because it is established with a secondary BS. However, the backup beam is not affected by the blocker crossing the primary beam path, as it uses an alternative beam providing a SNR that remains steady for the entire blockage duration.

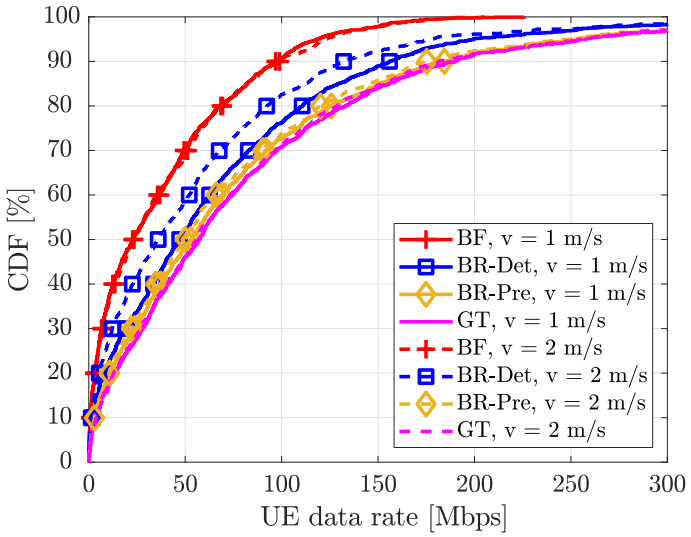
The BR-Pre method predicts the blockage state η time instants ahead of the time t as

(depicted with the symbol "x" in Fig. 5.4a). At the time $t - \eta$, if the blockage is predicted the BR-Pre method searches for an alternative beam pair and at time t, it switches to the backup beam. On the other hand, as we recall from Sec. 4.2.6, the BR-Det method switches to the backup beam with a delay β_1 as it searches for an alternative beam after detecting the blockage. In Figs. 5.4a and 5.4b, the BR-Pre method correctly predicts all blocked samples (TP cases), i.e. $\widehat{S}_l(t-\eta) = S_l(t) = 1$, meaning that no blockage states are miss detected, or, in other words, the models are not making FN errors. As a final point, the BR-Pre method predicts a good approximation of the number of successive blocked samples, regardless of the blocker speed. This indicates the ability to predict the blocker's duration as the speed changes without prior information about its value.

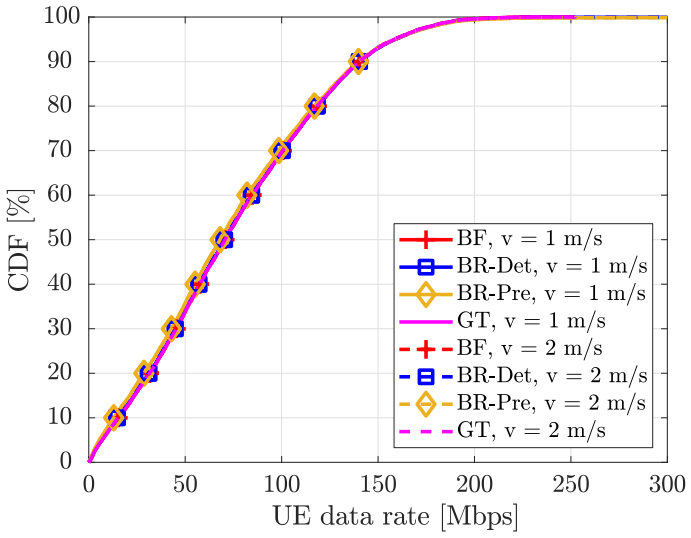
5.4.4 Data Rate Performance Comparison Between BR-Pre, BR-Det, BF and GT Methods

In these last three sections, we test the DNN models online into the system-level simulator, by generating an additional 10 network drops using different random seeds from the training dataset. We show the Cumulative Distribution Function (CDF) of the UEs data rate served by the BS-3 evaluated in the blocked and non-blocked intervals with the blocker moving at speeds 1 m/s and 2 m/s. We compare the performance between the beam fixed (BF) method that uses only the primary beam pair regardless of blockage events and the BR-Pre method, which uses the DNN models to make online predictions. We also show the results with the BR-Det method (see Sec. 4.2.6), and the Ground Truth (GT) method results representing an upper bound of BR-Pre method performance obtained when Eq. (5.4) has perfect knowledge of the future beam states.

The CDFs in Fig. 5.5a show that the BR-Pre method outperforms the BR-Det and BF methods during the blocked time instants, especially for the higher blocker speed. These results are explained as follows. We recall that the delay β_1 for detecting and switching the beam penalises the BR-Det data rate at the start of the blockage event. On the other hand, the BR-Pre method performance reaches 92% and 89% of the GT data rate due to the low number of FN errors that cause the BR-Pre method to use the primary beam during the blocked time instants. These indicate that the DNN models, when deployed online,



(a) CDF of UEs data rate for the blocked time instants $S_l(t) = 1$.

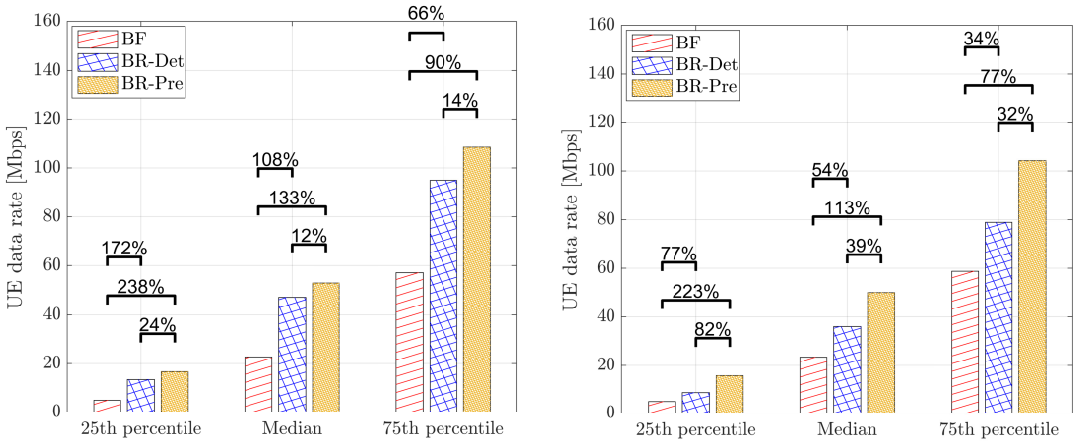


(b) CDF of UEs data rate for the non-blocked time instants $S_l(t) = 0$.

Figure 5.5: Data rate performance of the BS-3 beams for blocked and non-blocked time instants.

generalise well for both the blocker speeds. Overall, the DNN models wrong predictions have a minor impact on the data rate performance than the delay introduced for detecting the blockage and switching the beam. At a higher blocker speed, the beam switching happens more often, and the BR-Det data rate is penalised by the delay more times than with the blocker moving at 1 m/s, leading to a more significant difference to the BR-Pre results, as represented in Fig. 5.5a.

Additionally, Fig. 5.5b shows that in the non-blocked time instants, the BR-Pre data rates slightly deteriorate compared to the BR-Det data rate by 4% and 3% at the median



(a) UE data rate for blocker speed 1 m/s.

(b) UE data rate for blocker speed 2 m/s.

Figure 5.6: 25-th, 50-th (median) and 75-th percentiles of the data rate performance for BR-Pre, BR-Det and BF methods during blocked time instants. The percentages above bar charts indicate the data rate improvements by employing the BR-Pre and BR-Det methods.

of the two speeds CDFs, and by 6% and 5% at the 25-th percentile of the two speeds CDFs. These differences are due to the FP errors, which occur if the DNN models wrongly predict the beam state as blocked. Therefore, the BR-Pre method switches to the backup beam pair even if the primary beam pair is not blocked and the BR-Pre data rate is penalised because, during the non-blocked intervals, the primary beam has a larger SNR than the backup beam. These results show that during non-blocked time instants, the BR-Pre method performance difference to the BR-Det data rate, caused by the FP errors, are marginal.

5.4.5 Data Rate Performance for Worst Served UEs

In the remainder of this section, we take a deeper look at the data rate performance focusing on the lower part of the CDF presented in Fig. 5.5a, as the worst served UEs are most likely to suffer in blocked time instants to meet the requirement of the most demanding applications. Focusing on the 25th percentile results at blocker speed 1 m/s shown in Fig. 5.6a, the BR-Det and BR-Pre methods improve the BF data rate by 172% and 238% respectively and with BR-Pre outperforming BR-Det by 24%. Similarly, looking at 25th percentile results of Fig. 5.6b with the blocker moving at the speed of 2 m/s, the BR-Det and the BR-Pre methods improve the BF method data rate by 77% and 223%, respectively. The advantage of BR-Pre over BR-Det increases, and it is assessed at 82%. Note that the advantage offered by the BR-Pre method over the BR-Det one is more evident at the

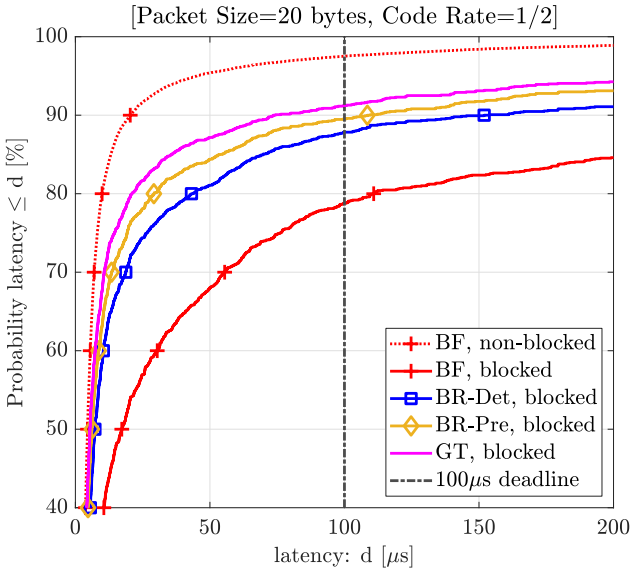
higher speed, as the reaction time of BR-Det does not scale with the higher frequency of the blockage events and their shorter duration, thus, by the time the method triggers the switch to the backup beam pair, the drop in the signal level caused by the blockage is either about to finish or already finished. In contrast, the BR-Pre data rate performance does not depend on the blocker speed, showing how predicting in advance the blockage occurrence is effective in indoor factory network deployments. Similar considerations apply for the median and the 75th percentiles of CDFs with blocker speed of 2 m/s showing 39% and 32% performance gains of our proposed prediction method over the detection one.

5.4.6 Air Interface Latency Performance Comparison between BF, BR-Det and BR-Pre Methods

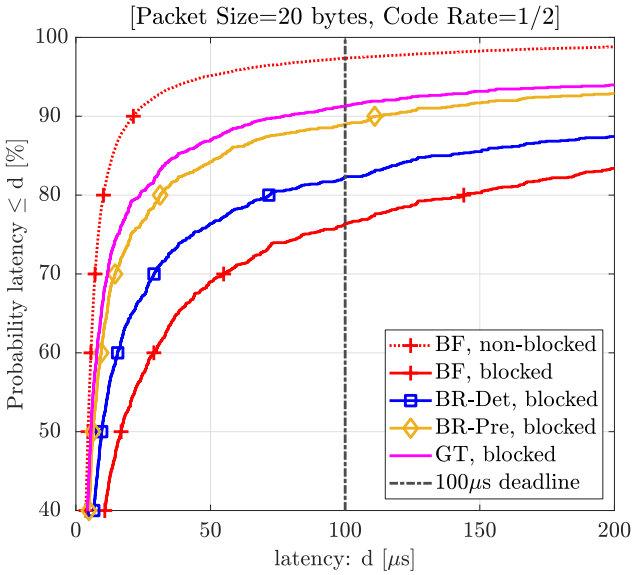
In this section, we compare the performance of the air interface latency between BF, BR-Pre and BR-Det methods. We consider transmitting a packet with a size of 20 bytes [70], and we account for the channel coding overhead given by a code rate of 1/2 [133]. The air interface latency can be expressed as $d_k(t) = \frac{P_{enc}}{r_k(t)}$, where $P_{enc} = 8 \times P_{inf} C_r$ is the encoded packet length in bits, C_r is the code rate and $r_k(t)$ is the data rate.

When computing the BR-Pre method latency, we neglect the delay due to the DNN models training, as we assume that the blockage conditions, e.g. blocker's movement and trajectory, repeat over time. In contrast, the blocker speed may change and the goal of the trained model is to make the predictions without knowing a prior the blocker speed. This use case is indicated for scenarios such as future industrial factories, where Automated Guided Vehicles (AGVs) or industrial robots may maintain similar movement patterns and follow a dedicated pathway in the factory corridor; however, they may change their speed depending on the carried load or production timeline.

In Fig. 5.7a, we compare the CDF of the latency values during blocked and non-blocked intervals for the different methods and blocker speed of 1 m/s. We also report the Physical layer (PHY) deadline of 100 μs, taken as a reference for the one way PHY transmission time [69]. The probability that the BF method's latency is lower than the PHY deadline is 97.5% during the non-blocked time instants, which indicates that the mmWave network without blockages would comply with the PHY deadline for most of the time. On the other hand,



(a) UE latency with blocker speed 1 m/s.



(b) UE latency with blocker speed 2 m/s.

Figure 5.7: CDF of UE latency for BF, BR-Det, Br-Pre and GT methods.

during the blocked intervals, this probability decreases significantly, and 78% of the latency values are lower than the PHY deadline of 100 μs. By adopting the BR-Det method, the probability increases to 88%, which is slightly lower than the BR-Pre method (90%) and the GT method (91%).

Additionally, in Fig. 5.7b, we show the same latency comparison between blocked and non-blocked intervals for the different methods, but we change the blocker speed to 2 m/s. The BF method probabilities achieve similar performance in Fig. 5.7a, i.e. 97% and 76% for

the non-blocked and blocked intervals. We maintain the identical blockage conditions the number of blockers and simulation duration stays the same. The BR-Pre method probability is 89% and stays close to the probability of the GT method (91%). This behaviour confirms what observed with the data rate, where the BR-Pre performance is similar to GT performance for both blocker speed used in the training dataset. Differently, the BR-Det method's latency probability decreases to 82%, which is lower than the value with the blocker speed of 1 m/s. This deviation confirms the trend observed with the data rate in Sec. 5.4.5, where the BR-Det method's performance deteriorates with the blocker moving at a higher speed because it is penalised by the higher frequency of the beam switching operations.

Overall, from the Fig. 5.7, we conclude that during the non-blocked intervals, the latency values' probability complies with the PHY deadline in most cases for both the blocker speeds. Instead, during the non-blocked intervals, this probability decreases significantly and only by adopting the BR-Pre method remain stable for the different blocker speeds.

5.5 Conclusion

In this chapter, we re-use existing beam measurement report messages from multiple UEs and BSs as input data for beam-specific DNN models that predict blockage events for multiple beams of the one exemplary BS. Our proposed multi-UE prediction-based method utilises the beam state predictions to control and trigger the beam recovery procedure ahead of the blockage events. This enables switching to a backup beam pair before the blockage disrupts the primary beam path. Thus, the prediction-based method allows for more stable signal quality and is more effective in significantly reducing the typical loss of data rate during the blockage event shown by commonly used methods based on the detection. The validation of the proposed method with online predictions shows how data rate performance achieves 92% and 89% of the GT data rate for the blockers speeds of 1 m/s and 2 m/s, respectively, confirming that the results are very close to the ideal case when the system has perfect knowledge of the future beam states. Moreover, the low number of FP errors rarely triggers the switching to the backup beam during non-blocked intervals and impacts the data rate performance marginally, losing only 4% and 3% to the BR-Det method median data rate. In addition, we show that the higher blocker speed penalises BR-Det method

as it is less effective in reacting to more frequent and shorter blockage events. In contrast, the BR-Pre data rate performance remains close to the GT. Finally, we found that for the worst served UE (at 25-th percentile of the CDF), the BR-Det improves the BF data rates by 172% and 77% for speeds 1 m/s and 2 m/s, while BR-Pre improves BF method data rates by 238% and 223% for speeds 1 m/s and 2 m/s. Overall, for the worst 25-th percentile UE, the proposed BR-Pre method improves the BR-Det method data rates by 24% at a blocker speed of 1 m/s and 82% at 2 m/s. The data rate improvements are also reflected in the results observed while analysing the air-interface latency. The probability that the latency value is lower than the PHY deadline in the blocked time instants increases with the BR-Pre compared to the BR-Det method for both the blocker speeds, especially when the blocker moves at the speed of 2 m/s.

6 Conclusion and Future Works

In this dissertation, we investigated novel network architectures and methods based on the combination of different 5-th Generation (5G) technologies (massive Multiple-Input-Multiple-Output (mMIMO), Ultra Dense Network (UDN), millimetre Wave (mmWave) and self-Backhauling (s-BH)) and emerging techniques such as Machine Learning (ML). In this chapter, we summarise the contributions made in our work; we give the main takeaways and discuss other important aspects that have not been covered and left for future studies.

6.1 Brief Summary of Results and Main Takeaways

In this section, we summarise the three main research questions and their answers, providing the main takeaways for each of them.

6.1.1 Research Question 1

Q-1: What are the end-to-end User Equipments (UEs) data rates performance when adopting the mMIMO-based s-BH network architecture? When to use mMIMO for s-BH? or when to use mMIMO to provide Direct Access (DA) service to UEs?

The s-BH technology aims to provide joint backhaul and access operations to facilitate the densification of the network. A compelling use case for s-BH is its application to Sub-6 GHz networks, where s-BH can be used outdoor to provide more flexibility to deploy Small Cells (SCs) while reducing the costs for the backhaul infrastructure. In addition, due to the scarcity of spectrum resources in Sub-6 GHz bands, we considered using mMIMO to serve multiple SCs in parallel utilising the same set of time-frequency resources.

In Chapter 3, we provided a comprehensive system-level simulation study to evaluate the interplay between access and backhaul and the resulting end-to-end UE data rates. Furthermore, we analysed the impact of different SCs deployment strategies while varying the time resource allocation between radio access and backhaul links. Finally, we compared the mMIMO-based s-BH network architecture to the mMIMO DA network architecture accounting for the effects of pilot reuse schemes, together with their associated overhead and contamination effects.

Main takeaways

Considering the same number of deployed SCs, the self-backhaul network architecture with the ad-hoc deployment of SCs outperforms the s-BH network architecture with the random deployment of SCs. This is because the access links achieve high signal-to-interference-and-noise ratios (SINRs) thanks to the gains given by the proximity between SCs and UEs. Furthermore, the results highlighted that selecting the appropriate resource partition between wireless access and backhaul links plays a key role in the end-to-end data rate performance. The optimal data rate can only be achieved throughout a carefully designed partition that trades access resources in favour of allocating more resources to the backhaul links.

Unlike mMIMO s-BH – where mMIMO-to-SC links are static, and thus channel acquisition is facilitated – mMIMO DA suffers more from pilot overhead and contamination. Compared to mMIMO DA solutions with pilot reuse scheme 3 and reuse scheme 1, ultra-dense SCs deployments supported by mMIMO s-BH provide rate improvements for cell-edge UEs that amount to 30% and a tenfold gain, respectively. This is because pilot contamination severely impacts the UEs performance at the cell-edge, and it is then better to serve the UEs with mMIMO s-BH architecture. Nevertheless, our analysis also showed that mMIMO DA outperforms s-BH above the median of the Cumulative Distribution Function (CDF) of the UE data rates, meaning that the access-backhaul bandwidth splitting limits the backhaul links' data rate of the s-BH architecture. Indeed, the losses due to this split outweigh the gains provided by the s-BH SCs' deployment.

These findings motivate us to investigate the potential benefit of using mmWave tech-

nology, which offers wide spectrum bandwidth to improve the backhaul's capacity that we found to be the main limitation for the s-BH aided mMIMO network architecture. Nevertheless, we observed that the self-backhauling network architecture requires having reliable wireless backhaul links, which needs to secure high capacity and stable backhauling connectivity to operate the access service. With mmWave systems, we expect that both backhaul and access links become sensitive to blockages, such as vehicles and other moving objects, making it challenging to guarantee a stable data rate. Therefore, we addressed the blockage problem on the mmWave access links in the following sections.

6.1.2 Research Question 2

Q-2: What are the effects of dynamic blockers when using mmWave for DA communications in an indoor scenario? How to improve the throughput stability by making the mmWave links more robust to blockage events?

5G networks require designing the network infrastructure to cope with new requirements, such as low latency and high reliability, which enable to operate, for instance, closed-loop control systems or allow real-time remote interaction with physical and virtual objects. As part of the key technologies for the 5G, mmWave communication can favour the deployment of Radio Access Networks (RANs), supporting very high data rates in the order of Gb/s, but it requires new solutions to cope with highly intermittent mmWave links due to the sensitivity to mobile blockages obstructing the signal paths. Therefore, one of the main challenges for designing mmWave systems is to improve the mmWave connections' robustness to the blockage events.

In Chapter 4, we studied a mmWave indoor network affected by dynamic blockages caused by mobile objects like humans and robots moving in the environment. Our study is more precise than others in the literature since our system-level simulator integrates a geometric blockage model identified as Blockage Model B, which enables us to simulate the smooth and continuous-time evolution of the blockage events. This is because the model accounts for the spatial consistency of the blockage effects, which is extremely important to effectively capture the temporal channel dynamics of the mmWave links and evaluate the beam recovery method's performance. To contrast the rapid changes in the signal strength

due to the blockage, we employed the BR-Det method to recover the blocked beam and we evaluate the network's data rate and latency performance varying the blocker's speed. Additionally, we quantified the blockage's impact on the probability of supporting a latency value lower than a realistic deadline for the Physical layer (PHY) data transmission.

Main takeaways

Our analysis showed that using a beam recovery method is of paramount importance as the blockage effect significantly impacts the data rate of the primary beam pair. The BR-Det method detects when the blocker enters the primary beam path and searches for a backup beam pair. Such a method increases the primary beam's median data rate by 80% when the blocker moves at a speed of 1 m/s. Moreover, the latency results showed that the blocker's presence increases the 95-th percentile latency value considerably above the 100 μs PHY deadline during the blocked intervals. Differently, when using the BR-Det method, the 95-th likely latency is lower than the deadline when the blockage probability is not larger than 20%. On the other hand, additional simulation results varying the blocker speed to 2 m/s revealed that the data rate gains of using the BR-Det method reduce. This behaviour is explained by higher blocker speed, which causes more frequent blockage events, increasing the beam switching frequency. Thus, making the delay due to the detection of the BR-Det method more penalising. These latest findings highlighted a potential problem for applying the BR-Det method to the mmWave indoor networks as it can be adversely affected by the blocker speed. To make the beam recovery operations more robust to such eventuality, we investigate alternative approaches.

6.1.3 Research Question 3

Q-3: How to avoid the delay for switching to the backup beam at every blockage event in indoor mmWave DA communications? How to design novel radio control methods by deploying ML models into the mmWave networks for making beam state predictions?

In mmWave communications, it is critical to study procedures that recover the link to limit the data rate losses happening during the blockage events. As shown in Chapter 4, reactive

methods yield a delay in the beam recovery operations as they switch to a new beam after detecting the blockage. In Chapter 5, we proposed a novel beam recovery method based on ML tools to predict the blockage events and switch to the backup beam before the link blockage occurs. Unlike the works presented in the literature focusing on the single-UE scenario, we tackled the more challenging multi-UEs network scenario requiring training a specific Deep Neural Network (DNN) model for each beam. Thus, we trained all the beam-specific DNN models of one Base Station (BS). Then, we deployed the DNN models online into the system-level simulator to measure the prediction-based method data rate and latency performance and compare them against the performance obtained with the detection-based method.

Main takeaways

We showed that the prediction-based method with DNN models deployed online achieves data rate performance close to the ground truth method, indicating the ability of the DNN models to generalise well for two different blocker speeds. Indeed, the DNN models online performance show a low number of false negative (FN) errors for both the blocker speeds used in the training dataset. During the blocked intervals, the prediction-based method guarantees higher signal level stability and up to 82% data rate improvement to the detection-based method when blockers move at a speed of 2 m/s. During the non-blocked intervals, the two methods achieve similar data rate performances. Additionally, also the air-interface latency shows improvements compared to the BR-Det method latency. During the blocked time instants, the probability that the latency value complies with the PHY deadline is higher with the beam recovery prediction method (BR-Pre) method than with the BR-Det method. This is explained by the fact that the detection's delay turns out to be detrimental for both the data rate and latency, especially at higher blocker speed. In these conditions, beam switching happens more frequently as the blocker moving faster in the indoor scenario increases the number of blockage events. On the other hand, the method based on predictions achieves performances that remain stable and close to the Ground Truth (GT) method performance.

The comparison between the method based on detection vs. the method based on

prediction revealed that:

- The beam recovery with detection is a simple yet practical approach. It can be applied to all the beams and does not require training; however, its performance degrades with the higher blocker speed when the beam-switching frequency increases. This represents a potential problem for applying the BR-Det beam recovery method to the mmWave indoor networks where the blocker speeds can be higher than an average pedestrian walking speed of ∼1 m/s. For example, Automated Guided Vehicles (AGVs) or mobile robots moving in an indoor factory move at speeds up to 8 m/s [57].
- The beam recovery method's performance with prediction remains close to the GT method performance for both the blocker speeds, making this approach more robust to the higher blocker speed. Nevertheless, this method requires employing beam-specific DNN models that may involve frequent updates if the blockage conditions change. This can be time-consuming to be handled manually by engineers because it involves repeating several steps, such as data collection, model training, and validation.

We conclude that, unlike the detection-based method that can be implemented once for all the beams, making the BR-Pre recovery method feasible for practical network deployments necessitate to automate the data collection, training, and validation procedures through a ML workflow, which can make these operations repeatable for each beam and easy to be updated when the blockage conditions change.

6.1.4 General Research Question

How will the combination of 5G technologies and other emerging techniques impact the design of novel network architectures and functionalities to deliver seamless wireless connectivity that supports multiple 5G services in the years to come?

This dissertation showed two examples in which the combination of 5G technologies and new enabling technologies provides advanced network designs that improve the network data rate and air interface latency performance. The two key takeaways of this dissertation can be summarised as follow:

- In the *outdoor* scenario, Sub-6 GHz mMIMO systems, besides providing DA service to their UEs, can serve a massive number of SCs that connect utilising the s-BH technology to the mMIMO-BS. The s-BH technology guarantees the flexibility to dynamically reposition the SCs to reach locations close to the UEs and improve the UEs data rate at the cell edge, where pilot contamination affects the performance of DA connections.
- In the *indoor* scenario, a beam recovery method based on ML predictions mitigates the blockages caused on mmWave links by the movement of objects in the environment. The BR-Pre method provides a more stable mmWave link data rate than conventional beam recovery methods based on detection and reduces the air-interface latency thanks to the capability of the DNN models to generalise well for different blocker speeds.

We conclude that a key aspect of deploying these networks and methods will be developing automated procedures to facilitate wireless network configuration and optimisation. Software-based components and ML tools can be the two main directions to take as they can facilitate such automation. For instance, the Open-Radio Access Network (O-RAN) provides all the network architecture elements to develop software-defined components and implement a ML workflow for optimising its performance. In the next section, we show an example of a ML workflow for implementing the prediction-based method.

6.2 Future Works

In this section, we discuss important topics that have not been covered in this dissertation, and we have left for future studies.

6.2.1 Hybrid Beamforming

We recall that the hybrid beamforming architecture combines a digital precoder stage with multiple Radio Frequency (RF) chains in the baseband processing and analog beamforming in the RF domain [84, 140]. First, this type of beamforming architecture adds more Degrees-of-Freedom (DoF) to the multiple-input-multiple-output (MIMO) operations as multiple

data transmission streams can be transmitted simultaneously utilising various beamforming directions at the same time. Secondly, the digital stage in the baseband processing enables the beamforming patterns to be adapted through the precoding coefficients while using the analog beamforming stage to modify the steering directions and improve the beamforming gain. Furthermore, novel RF designs consider low Analog-to-Digital converter (ADC) resolution at the mixed analog-digital device and limits the total power consumption of the ADC. Nevertheless, a lower ADC resolution introduces hardware non-linearities, which add a quantisation error to the baseband signal representation [141].

The utilisation of the hybrid beamforming architecture at the network level would improve several aspects of the mmWave system design discussed in this dissertation: *i)* enabling the utilisation of mmWave system for wireless backhauling to serve multiple s-BHs SCs in parallel, *ii)* providing more flexible beam switching operations and *iii)* improving the mmWave link robustness to blockage as the UE can be served from multiple directions simultaneously. However, the utilisation of hybrid beamforming in mmWave cellular deployments would introduce several new challenges, such as the large overhead due to the channel estimation for the configuration of the digital precoder and analog beamforming.

Overall, all these aspects require a detailed system-level analysis that accounts for the hardware non-linearities and evaluates the network performance tradeoffs between utilising fully-analog or hybrid architectures.

6.2.2 Ultra-Reliable Low-Latency Communication (URLLC)

This dissertation focused on the data rate and the air interface latency of the 5G communication systems. An additional aspect that characterises URLLC is the service reliability requirement. The reliability quantifies the probability of transmitting and successfully decoding a packet of P_{inf} Bytes by assuming a fixed deadline, e.g. 1 ms end-to-end UE plane latency [142]. This probability is computed by measuring the block error rate (BLER) after the decoders. Thus, depending on the type of application, the target BLER may decrease from 10^{-5} to 10^{-9}, which correspond to the reliability requirements of five-nines and nine-nines, respectively. Vertical industries such as Industry 4.0, automotive, health broadcasting and media, support several use cases with reliability requirements of five-nines (99.999%),

e.g. gaming and Virtual Reality (VR)-Augmented Reality (AR), or even nine-nines in the most conservative cases, e.g. industrial robots. Additionally, reliability requirements may be evaluated employing a PHY layer abstraction model that estimates the PHY transmission channel coding performance [143]. We left this analysis for future studies as it changes several assumptions of our system model, such as traffic conditions and the number of devices accessing the channel, to name a few. All these modifications of the system model require conducting a specific and detailed analysis.

6.2.3 ML Prediction for Non-uniform Blockage Movements

In this section, we discuss possible extensions of the DNN model to predict the blockage on a beam. We recall that the model proposed in Chapter 5 supports the uniform blockage movement with different speeds. Nevertheless, the ML model could be extended to support non-uniform object movements, e.g. with a blocker accelerating or decelerating. In the former case, a static blocker may start moving and, for instance, increases its speed v from 0 to 3 m/s, consequently blocking some of the beams on its trajectory. In the latter case, a dynamic blocker may decrease the speed v from 3 to 0 m/s without causing further blockage events. In both cases, the blockage model proposed in Chapter 5 may provide inaccurate predictions since the data used during training present a uniform speeds. Moreover, we assumed constant sampling time $\Delta_t = 200$ ms over the observation window $\epsilon = 2$ s and we fixed the prediction window to $\eta = 400$ ms, which may become a limitation with the blocker accelerating.

Therefore, the first challenge is to train new DNN models adapting the set of parameters like the prediction window η, the observation window ϵ and the sampling time Δ_t to support non-uniform blockage movements. For the scenario of blocker accelerating, we expect that the signal levels of the set of beams used as input of the DNN model show faster variations within the observation window ϵ as the blockage speed increases over time. Therefore, we expect the need to reduce the sampling time Δ_t to update the DNN model's input more frequently and maintain a constant spatial resolution. Also, the prediction window η may require to be reduced proportionally to the sampling time adaptation. On the other hand, when the blocker decelerates, we expect that the DNN model provides accurate predictions,

even with a larger sampling time step within the observation window ϵ and we may increase the prediction window η accordingly.

As the prediction parameters may change, we expect that both types of movements require the training of specific DNN models. Thereby, a second challenge is to select the output from one of the models when deploying the models online.

6.2.4 ML Prediction Workflow in Open-RAN

In this section, we provide an example of ML workflow for implementing the prediction-based method, according to the standard directives of the O-RAN architecture [144]. Then, we explain the ML workflow, including training, validation and online deployment, considering the various O-RAN architecture components designed for hosting these functions. Starting from the bottom of Fig. 6.1, we describe the different steps of the O-RAN ML workflow to enable the BR-Pre as follows:

1. The open front-haul interface connects the mmWave antenna arrays to a centralised server, which hosts the baseband functions of the multiple BSs.
2. The BSs exchange the beam measurements data through a central server. The O1 interface performs data collection and sends the beam measurements to the Non-Real-Time RAN Intelligent Controller (Non-RT RIC).
3. The Non-RT RIC supports the offline training of ML models that are later collected in a dedicated catalog.
4. The ML models are selected based on the performance metrics such as F1-score (see Appendix B) and are deployed in the near-RT RIC.
5. A ML-based external application (xApp) – designed to optimise the RAN functions – collects measurements data from the central server through the E2 interface and makes predictions about the blockage state of the beams.
6. The xApps forward the ML model output to a control loop (Loop2), which updates with a near real-time periodicity, e.g. 200 ms, the beams' blockage states.
7. The BR-Pre method initiates the pro-active beam switching accordingly to the beam state predictions, as shown before in Sec. 5.2.2.

The network performance monitored real-time, may indicate whether the ML models

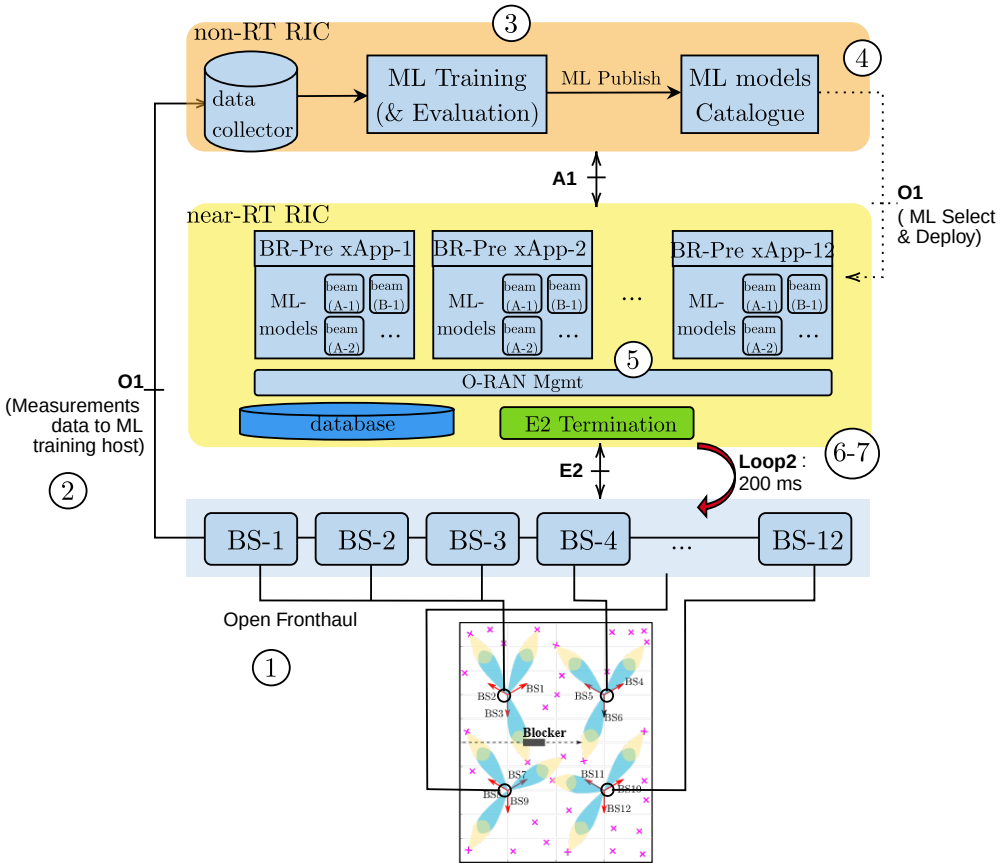


Figure 6.1: Open-RAN Implementation Example.

require to be updated and re-trained. In that case, the ML workflow iterates steps $1-4$ of the procedure described above.

To conclude, we have shown that it is feasible to use the ML workflow of the O-RAN architecture for implementing the prediction-based method. Thus, the completion of this workflow can enable the realise automated procedure to optimise the network performance. Nevertheless, each phase of the ML workflow requires to be properly engineered, and we leave this analysis for further studies.

7 Appendix

A Rate Coverage Probability Expression for Small Cell Access Transmission

In the following, we derive the expression used to evaluate the rate coverage probability. The probability that the access SIR is greater than a threshold $\gamma_a = 2^{R_{th}\mu_l} - 1$, which depends on the minimum target rate R_{th}, is expressed as [117]

$$\Pr\left[\mathrm{SIR}^{\mathrm{A}}(x) > \gamma_a\right] = \Pr\left[\frac{P_l^{\mathrm{dl}} G_b |g|^2 \beta^{\mathrm{L}}(x)}{I_{\mathrm{agg}}} > \gamma_a\right] \overset{(a)}{=} \mathscr{L}_{I_{\mathrm{agg}}}\left(\frac{\gamma_a}{P_l^{\mathrm{dl}} G_b \beta^{\mathrm{L}}(x)}\right), \tag{7.1}$$

where (a) follows from [30, Eq. (54)] neglecting the thermal noise as the propagation in sub-6 GHz bands is interference-limited and $\mathscr{L}_{I_{\mathrm{agg}}}(s)$ represents the Laplace transform of I_{agg}, and is defined according to [30, 117] as follows

$$\begin{aligned}\mathscr{L}_{I_{\mathrm{agg}}}(s) = &\exp\left(-2\pi\widetilde{\lambda}_b \int_x^{+\infty} \frac{\Pr^{\mathrm{L}}(u)\, u}{1 + \left(s P_l^{\mathrm{dl}} G_b \beta^{\mathrm{L}}(u)\right)^{-1}} du\right) \\ &\times \exp\left(-2\pi\widetilde{\lambda}_b \int_{x_1}^{+\infty} \frac{\left[1 - \Pr^{\mathrm{L}}(u)\right] u}{1 + \left(s P_l^{\mathrm{dl}} G_b \beta^{\mathrm{NL}}(u)\right)^{-1}} du\right),\end{aligned} \tag{7.2}$$

where $x_1 = \left(\frac{A^{\mathrm{NL}}}{A^{\mathrm{L}}}\right)^{\eta^{\mathrm{NL}-1}} x^{\frac{\eta^{\mathrm{L}}}{\eta^{\mathrm{NL}}}}$ [117].

B ML Performance Measures

In the following we provide a detailed description of the metrics used in Sec. 5.4.2 to validated the Deep Neural Network (DNN) models and we show the confusion matrix, which cross-tabulates the outcomes of the DNN model and helps to evaluate the binary classifier's performance.

Table 7.1: Confusion matrix used to evaluate the output of a beam state classifier.

		Predicted outcome $\widehat{Y}_n = 0$ (non-blocked)	$\widehat{Y}_n = 1$ (blocked)	Total
Ground truth	$Y_n = 0$ (non-blocked)	TN	FP	TN + FP
	$Y_n = 1$ (blocked)	FN	TP	FN + TP
	Total	TN + FN	FP + TP	N

As shown in Table 7.1, the main diagonal elements of the confusion matrix, i.e. true positive (TP) and true negative (TN), represent the cases when the classifier correctly predicts the state blocked and non-blocked, respectively. On the other hand, the off-diagonal elements, i.e. false positive (FP) and false negative (FN), refer to two different errors. In the first case, the classifier wrongly predicts the state blocked when the actual state is non-blocked. While in the latter, the classifier incorrectly predicts the state non-blocked when the actual state is blocked.

The accuracy indicates the number of predicted samples (TP+TN) that the model has correctly classified, and can be expressed as

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{7.3}$$

This measure is not appropriate for a reliable indication of the model performance when the data present unbalanced classes because it does not differentiate between the errors committed on the rare class and the errors committed on the dominant class. Consequently, the model may achieve very high accuracy even if it does not classify the rare class samples. To avoid this type of error in the analysis, we opted to evaluate the classifier performances considering other measures such as Precision, Recall and F1-score that are more indicated

with unbalanced classes [138]. We provide their definition as follows:

- Recall indicates the proportion of actual blocked samples (positive cases), which are predicted correctly and can be expressed as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{7.4}$$

- Precision indicates the proportion of actual blocked samples among all predicted blocked samples and can be expressed as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{7.5}$$

- The F1-score combines through the harmonic mean Precision and Recall and can be expressed as

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7.6}$$

Bibliography

[1] Qualcom, "The 1000x mobile data challenge," Webinar, Aug. 2012.

[2] D. López-Pérez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2078–2101, Fourthquarter 2015.

[3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Dec. 2013.

[4] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, Jun. 2014.

[5] Wikipedia, "List of 5G NR networks," available at https://en.wikipedia.org/wiki/List_of_5G_NR_networks, Last accessed: May 2021.

[6] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[7] H. Willebrand and B. Ghuman, "Fiber optics without fiber," *IEEE Spectrum*, vol. 38, no. 8, pp. 40–45, 2001.

[8] 3GPP Technical Report (TR) 38.874, "NR; Study on integrated access and backhaul," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Oct. 2019.

[9] A. Aijaz, "Private 5G: The future of industrial wireless," *IEEE Industrial Electronics Magazine*, vol. 14, no. 4, pp. 136–145, Dec. 2020.

[10] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[11] S. H. A. Shah, M. Sharma, and S. Rangan, "LSTM-based multi-link prediction for mmWave and Sub-THz wireless systems," in *2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[12] M. Alrabeiah and A. Alkhateeb, "Deep learning for mmWave beam and blockage prediction using Sub-6 GHz channels," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5504–5518, Sep. 2020.

[13] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: Model-based, AI-based, or both?" *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.

[14] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 4, pp. 3072–3108, Fourthquarter. 2019.

[15] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[16] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019.

[17] C. Madapatha, B. Makki, C. Fang, O. Teyeb, E. Dahlman, M.-S. Alouini, and T. Svensson, "On integrated access and backhaul networks: Current status and potentials," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1374–1389, 2020.

[18] Nokia Bell Labs, "F-Cell technology from Nokia Bell Labs revolutionizes small cell deployment by cutting wires, costs and time," Oct. 2016, https://www.nokia.com/about-us/news/releases/2016/10/03/f-cell-technology-from-nokia-bell-labs-revolutionizes-small-cell-deployment-by-cutting-wires-costs-and-time/.

[19] Terragraph by Facebook, "Terragraph: Solving the urban bandwidth challenge," 2018, https://terragraph.com/terragraph.

[20] A. Bonfante, L. Galati Giordano, D. López-Pérez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, M. M. Butt, and N. Marchetti, "5G massive MIMO architectures: Self-backhauled small cells versus direct access," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10 003–10 017, Oct. 2019.

[21] A. Bonfante, L. Galati Giordano, I. Macaluso, and N. Marchetti, "Performance of predictive indoor mmWave networks with dynamic blockers," *arXiv preprint arXiv:2104.04623*, Apr. 2021.

[22] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[23] H. Claussen, D. López-Pérez, L. Ho, R. Razavi, and S. Kucera, "Small cell networks: Deployment, management, and optimization," *Wiley-IEEE Press*, 2018.

[24] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced," *IEEE Wireless Communications*, vol. 17, no. 3, pp. 26–34, 2010.

[25] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A flexible 5G frame structure design for frequency-division duplex cases," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 53–59, 2016.

[26] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.

[27] L. Zhang, A. Ijaz, P. Xiao, M. M. Molu, and R. Tafazolli, "Filtered OFDM systems, algorithms, and performance analysis for 5G and beyond," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1205–1218, 2018.

[28] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, "Fundamentals of massive MIMO," *Cambridge University Press*, 2016.

[29] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 4, pp. 2017–2046, 2015.

[30] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin, "Performance impact of LoS and NLoS transmissions in dense cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 2365–2380, Mar. 2016.

[31] D. López-Pérez and M. Ding, "Toward ultradense small cell networks: A brief history on the theoretical analysis of dense wireless networks," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–11, May 2019.

[32] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," *IEEE Network*, vol. 28, no. 6, pp. 6–11, Nov. 2014.

[33] M. Jaber, F. J. Lopez-Martinez, M. A. Imran, A. Sutton, A. Tukmanov, and R. Tafazolli, "Wireless backhaul: Performance modeling and impact on user association for 5g," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3095–3110, May 2018.

[34] U. Siddique, H. Tabassum, E. Hossain, and D. I. Kim, "Wireless backhauling of 5G small cells: challenges and solution approaches," *IEEE Wireless Communications*, vol. 22, no. 5, pp. 22–31, Oct. 2015.

[35] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Communications*, vol. 9, no. 11, pp. 3590–3600, Oct. 2010.

[36] B. Li, D. Zhu, and P. Liang, "Small cell in-band wireless backhaul in massive MIMO systems: A cooperation of next-generation techniques," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 7057–7069, Dec. 2015.

[37] H. H. Yang, G. Geraci, and T. Q. S. Quek, "Energy-efficient design of MIMO heterogeneous networks with wireless backhaul," *IEEE Transactions on Wireless Communications*, vol. 15, no. 7, pp. 4914–4927, Jul. 2016.

[38] G. Geraci, A. Garcia-Rodriguez, D. López-Pérez, A. Bonfante, L. Galati Giordano, and H. Claussen, "Operating massive MIMO in unlicensed bands for enhanced coexistence and spatial reuse," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1282–1293, Jun. 2017.

[39] G. Geraci, A. Garcia-Rodriguez, D. López-Pérez, A. Bonfante, L. G. Giordano, and H. Claussen, "Enhancing coexistence in the unlicensed band with massive MIMO," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.

[40] A. Garcia-Rodriguez, G. Geraci, L. G. Giordano, A. Bonfante, M. Ding, and D. López-Pérez, "Massive MIMO unlicensed: A new approach to dynamic spectrum access," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 186–192, Jun. 2018.

[41] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *2014 IEEE Global Telecommunications Conference Workshops (GLOBECOM Workshops)*, 2014, pp. 287–293.

[42] FCC, "Application for original equipment certification under FCC ID: VBNAAHE-01 for Nokia AirScale MAA 64T64R 128AE B41 120W AAHE, operating in the Band 41," Jul. 2018, https://fcc.report/FCC-ID/VBNAAHE-01.

[43] L. Guan, L. G. Giordano, and A. Bonfante, "A flexible HW and SW co-operated baseband research platform for massive MIMO system," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–7.

[44] T. L. Marzetta and B. M. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1268–1278, Mar. 2006.

[45] L. Galati Giordano, L. Campanalonga, D. López-Pérez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, and M. Magarini, "Uplink sounding reference signal coordination to combat pilot contamination in 5G massive MIMO," in *Proc. IEEE Wireless Commun. Networking Conference (WCNC)*, Apr. 2018, pp. 1–6.

[46] 3GPP Technical Specification (TS) 38.101-2, "NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone," 3rd Generation Partnership Project (3GPP), Version 16.3.1, Apr. 2020.

[47] Y. Ghasempour, C. R. C. M. da Silva, C. Cordeiro, and E. W. Knightly, "IEEE 802.11ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 186–192, Dec. 2017.

[48] ITU-R, "Radio-Frequency Channel and Block Arrangements for Fixed Wireless Systems Operating in the 71-76 and 81-86 GHz Bands," Recommendation ITU-R F.2006, Mar. 2012.

[49] T. Baykas, C. Sum, Z. Lan, J. Wang, M. A. Rahman, H. Harada, and S. Kato, "IEEE 802.15.3c: the first IEEE wireless standard for data rates over 1 Gb/s," *IEEE Communications Magazine*, vol. 49, no. 7, pp. 114–121, Jun. 2011.

[50] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11ad: directional 60 GHz communication for multi-gigabit-per-second Wi-Fi [invited paper]," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 132–141, Dec. 2014.

[51] R. Ford, M. Zhang, M. Mezzavilla, S. Dutta, S. Rangan, and M. Zorzi, "Achieving ultra-low latency in 5G millimeter wave cellular networks," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 196–203, Mar. 2017.

[52] G. Yang, M. Xiao, and H. V. Poor, "Low-latency millimeter-wave communications: Traffic dispersion or network densification?" *IEEE Transactions on Communications*, pp. 1–1, Mar. 2018.

[53] S. Kutty and D. Sen, "Beamforming for millimeter wave communications: An inclusive survey," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 2, pp. 949–973, Secondquarter 2016.

[54] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 1, pp. 173–196, Firstquarter 2019.

[55] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[56] R. J. Weiler, M. Peter, W. Keusgen, and M. Wisotzki, "Measuring the busy urban 60 GHz outdoor access radio channel," in *2014 IEEE International Conference on Ultra-WideBand (ICUWB)*, 2014, pp. 166–170.

[57] 3GPP Technical Report (TR) 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Jan. 2020.

[58] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, "Small-scale, local area, and transitional millimeter wave propagation for 5G communications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6474–6490, Dec. 2017.

[59] M. Gapeyenko, A. Samuylov, M. Gerasimenko, D. Moltchanov, S. Singh, M. R. Akdeniz, E. Aryafar, N. Himayat, S. Andreev, and Y. Koucheryavy, "On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 124–10 138, Nov. 2017.

[60] J. Park, J. Lee, J. Liang, K. Kim, K. Lee, and M. Kim, "Millimeter wave vehicular blockage characteristics based on 28 GHz measurements," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[61] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," *Proceedings of the IEEE*, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.

[62] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 436–453, Feb. 2016.

[63] Junyi Wang, Zhou Lan, Chang-woo Pyo, T. Baykas, Chin-sean Sum, M. A. Rahman, Jing Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, pp. 1390–1399, Sep. 2009.

[64] M. N. Kulkarni, J. G. Andrews, and A. Ghosh, "Performance of dynamic and static TDD in self-backhauled millimeter wave cellular networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 10, pp. 6460–6478, Oct. 2017.

[65] S. Singh, M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "Tractable model for rate in self-backhauled millimeter wave cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2196–2211, Oct. 2015.

[66] R. Gupta and S. Kalyanasundaram, "Resource allocation for self-backhauled networks with half-duplex small cells," in *2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 198–204.

[67] T. M. Nguyen, A. Yadav, W. Ajib, and C. Assi, "Resource allocation in two-tier wireless backhaul heterogeneous networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6690–6704, Oct. 2016.

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

[69] G. P. Fettweis, "The tactile internet: Applications and challenges," *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 64–70, Mar. 2014.

[70] 3GPP Technical Report (TR) 22.804, "Study on Communication for Automation in Vertical Domains (Release 16)," 3rd Generation Partnership Project (3GPP), Version 16.3.0, Jul. 2020.

[71] J. Kibiłda, A. B. MacKenzie, M. J. Abdel-Rahman, S. K. Yoo, L. G. Giordano, S. L. Cotton, N. Marchetti, W. Saad, W. G. Scanlon, A. Garcia-Rodriguez, D. López-Pérez, H. Claussen, and L. A. DaSilva, "Indoor millimeter-wave systems: Design and performance evaluation," *Proceedings of the IEEE*, vol. 108, no. 6, pp. 923–944, Jun. 2020.

[72] Ofcom, "Enabling wireless innovation through local licensing," Guidance document, Jul. 2019.

[73] A. Rao, "ML/AI-based automated assurance is critical for the success of 5G," Analysys Mason, White Paper, Jan. 2020.

[74] T. L. Marzetta, O. Blume, P. Rulikowski, S. Maier, A. Pascht, and T. Klein, "Frequency division duplex FDD Massive MIMO backhaul for repeater small cells," Granted Patent US 9 768 983 B2, Sep., 2017.

[75] S. Wesemann and T. L. Marzetta, "Channel training for analog FDD repeaters: Optimal estimators and Cramér-Rao bounds," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6158–6170, Sep. 2017.

[76] T. Danford, O. Filiz, J. Huang, B. Karrer, M. Paluri, G. Pang, V. Ponnampalam, N. S. Moses, and B. Tezel, "End-to-end planning of fixed millimeter-wave networks," *arXiv preprint arXiv:1705.07249*, Mar. 2017.

[77] H. Tabassum, A. H. Sakr, and E. Hossain, "Analysis of massive MIMO-enabled downlink wireless backhauling for full-duplex small cells," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2354–2369, Jun. 2016.

[78] C. Slezak, M. Zhang, M. Mezzavilla, and S. Rangan, "Understanding end-to-end effects of channel dynamics in millimeter wave 5G new radio," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018, pp. 1–5.

[79] I. K. Jain, R. Kumar, and S. S. Panwar, "The impact of mobile blockers on millimeter wave cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, pp. 854–868, Apr. 2019.

[80] W. Yang and K. Liu, "Blockage effect and beam cooperation in indoor hotspot based on 3GPP NR blockage model," in *2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[81] Y. M. Tsang and A. S. Y. Poon, "Detecting human blockage and device movement in mmWave communication system," in *2011 IEEE Global Telecommunications Conference (GLOBECOM) 2011*, 2011, pp. 1–6.

[82] M. Jasim, M. Ababneh, N. Siasi, and N. Ghani, "Hybrid beamforming for link recovery in millimeter wave communications," in *2018 IEEE 19th Wireless and Microwave Technology Conference (WAMICON)*, 2018, pp. 1–4.

[83] Y. Li, F. Baccelli, J. G. Andrews, and J. C. Zhang, "Directional cell search delay analysis for cellular networks with static users," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 4318–4332, Sep. 2018.

[84] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[85] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting, "Millimeter wave beam alignment: Large deviations analysis and design insights," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1619–1631, Jul. 2017.

[86] 3GPP Technical Specification (TS) 38.321, "NR; Medium Access Control (MAC) protocol specification," 3rd Generation Partnership Project (3GPP), Version 16.0.0, Apr. 2020.

[87] M. Zarifneshat, L. Xiao, and J. Tang, "Learning-based blockage prediction for robust links in dynamic millimeter wave networks," in *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, June 2019, pp. 1–9.

[88] A. Alkhateeb, I. Beltagy, and S. Alex, "Machine learning for reliable mmWave systems: blockage prediction and proactive handoff," in *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018, pp. 1055–1059.

[89] Z. Ali, A. Duel-Hallen, and H. Hallen, "Early warning of mmWave signal blockage and AoA transition using sub-6 GHz observations," *IEEE Communications Letters*, vol. 24, no. 1, pp. 207–211, Jan. 2020.

[90] T. Nishio, Y. Koda, J. Park, M. Bennis, and K. Doppler, "When wireless communications meet computer vision in beyond 5G," *arXiv preprint arXiv:2010.06188*, Oct. 2020.

[91] T. Nishio, H. Okamoto, K. Nakashima, Y. Koda, K. Yamamoto, M. Morikura, Y. Asai, and R. Miyatake, "Proactive received power prediction using machine learning and depth images for mmWave networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2413–2427, Nov. 2019.

[92] Y. Koda, K. Nakashima, K. Yamamoto, T. Nishio, and M. Morikura, "Handover management for mmWave networks with proactive performance prediction using camera images and deep reinforcement learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 802–816, Jun. 2020.

[93] L. Simić, J. Arnold, M. Petrova, and P. Mähänen, "RadMAC: Radar-enabled link obstruction avoidance for agile mm-wave beamsteering," in *Proceedings of the 3rd Workshop on Hot Topics in Wireless*, 2016, pp. 61–65.

[94] 3GPP Technical Specification (TS) 38.300, "NR; Overall description; Stage-2," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Apr. 2020.

[95] A. Bonfante, L. G. Giordano, D. López-Pérez, A. Garcia-Rodriguez, G. Geraci, P. Baracca, M. M. Butt, M. Dzaferagic, and N. Marchetti, "Performance of massive MIMO self-backhauling for ultra-dense small cell deployments," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec 2018, pp. 1–7.

[96] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, Jul. 2015.

[97] 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, Oct. 2015.

[98] 3GPP Technical Report (TR) 36.814, "Further advancements for E-UTRA physical layer aspects," 3rd Generation Partnership Project (3GPP), Version 9.2.0, Mar. 2017.

[99] M. Feng, S. Mao, and T. Jiang, "Joint frame design, resource allocation and user association for massive MIMO heterogeneous networks with wireless backhaul," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1937–1950, Mar. 2018.

[100] C. Liu, M. Ding, C. Ma, Q. Li, Z. Lin, and Y. Liang, "Performance analysis for practical unmanned aerial vehicle networks with LoS/NLoS transmissions," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2018, pp. 1–6.

[101] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving "Massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.

[102] 3GPP Technical Report (TR) 25.996, "Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," 3rd Generation Partnership Project (3GPP), Version 16.0.0, Jul. 2020.

[103] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering, and keyhole," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

[104] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1924–1934, Jul. 2007.

[105] A. Molisch, "Wireless communications," *John Wiley and Sons*, 2010.

[106] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, vol. 5, no. 1, pp. 10–22, Feb. 1998.

[107] X. Zhu, Z. Wang, C. Qian, L. Dai, J. Chen, S. Chen, and L. Hanzo, "Soft pilot reuse and multicell block diagonalization precoding for massive MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3285–3298, May 2016.

[108] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[109] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[110] M. Haenggi, "Stochastic geometry for wireless networks," *Cambridge University Press*, 2012.

[111] C. Li, J. Zhang, and K. B. Letaief, "Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2505–2517, May 2014.

[112] S. Lee and K. Huang, "Coverage and economy of cellular networks with many base stations," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1038–1040, Jul. 2012.

[113] M. Minelli, M. Ma, M. Coupechoux, J. M. Kelif, M. Sigelle, and P. Godlewski, "Optimal relay placement in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 998–1009, Feb. 2014.

[114] J. M. Kelif, M. Coupechoux, and P. Godlewski, "A fluid model for performance analysis in cellular networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, no. 435189, Aug. 2010.

[115] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, "Statistical analysis and modelling of spatial point patterns," *John Wiley and Sons*, 2008.

[116] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 156–163, Nov. 2010.

[117] C. Galiotto, N. K. Pratas, N. Marchetti, and L. Doyle, "A stochastic geometry framework for LoS/NLoS propagation in dense small cell networks," in *2015 IEEE International Conference on Communications (ICC)*, Jun. 2015, pp. 2851–2856.

[118] S. Singh, H. S. Dhillon, and J. G. Andrews, "Offloading in heterogeneous networks: Modeling, analysis, and design insights," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2484–2497, May 2013.

[119] G. George, A. Lozano, and M. Haenggi, "Distribution of the number of users per base station in cellular networks," *IEEE Wireless Communications Letters*, vol. 8, no. 2, pp. 520–523, Apr. 2019.

[120] J. Kibiłda and G. de Veciana, "Dynamic network densification: Overcoming spatio-temporal variability in wireless traffic," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2018, pp. 1–6.

[121] T. Bai and R. W. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Transactions on Communications*, vol. 14, no. 2, pp. 1100–1114, Oct. 2015.

[122] G. R. MacCartney, T. S. Rappaport, and S. Rangan, "Rapid fading due to human blockage in pedestrian crowds at 5G millimeter-wave frequencies," in *2017 IEEE Global Communications Conference (GLOBECOM)*, 2017, pp. 1–7.

[123] 3GPP Technical Specification (TS) 38.211, "NR; Physical channels and modulation," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Apr. 2020.

[124] 3GPP Technical Report (TR) 38.912, "Study on new radio (nr) access technology," Version 15.0.0, Jun. 2018.

[125] S. Dutta, M. Mezzavilla, R. Ford, M. Zhang, S. Rangan, and M. Zorzi, "MAC layer frame design for millimeter wave cellular system," in *2016 European Conference on Networks and Communications (EuCNC)*, 2016, pp. 117–121.

[126] 3GPP Technical Specification RT-170019, "Summary of Email Discussion [ITU-R AH 01] Calibration for Self-Evaluation," 3rd Generation Partnership Project (3GPP), 3GPP TSG RAN ITU-R ad hoc, Dec. 2017.

[127] Q. Nadeem, A. Kammoun, and M. Alouini, "Elevation beamforming with full dimension MIMO architectures in 5G systems: A tutorial," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 4, pp. 3238–3273, Fourthquarter 2019.

[128] A. Kammoun, H. Khanfir, Z. Altman, M. Debbah, and M. Kamoun, "Preliminary results on 3D channel modeling: From theory to standardization," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1219–1229, 2014.

[129] J. Kunisch and J. Pamp, "Ultra-wideband double vertical knife-edge model for obstruction of a ray by a person," in *2008 IEEE International Conference on Ultra-Wideband*, vol. 2, Sep. 2008, pp. 17–20.

[130] 3GPP Technical Report (TR) 38.802, "Study on new radio access technology physical layer aspects," Version 14.2.0, Sep. 2017.

[131] 3GPP Technical Specification (TS) 38.213, "NR; Physical layer procedures for control," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Apr. 2020.

[132] M. Tayyab, X. Gelabert, and R. Jäntti, "A survey on handover management: From LTE to NR," *IEEE Access*, vol. 7, pp. 118 907–118 930, Aug. 2019.

[133] D. Hui, S. Sandberg, Y. Blankenship, M. Andersson, and L. Grosjean, "Channel coding in 5G New Radio: A tutorial overview and performance comparison with 4G LTE," *IEEE Vehicular Technology Magazine*, vol. 13, no. 4, pp. 60–69, Oct. 2018.

[134] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, Nov. 2018.

[135] Y. Wang, A. Klautau, M. Ribero, A. C. K. Soong, and R. W. Heath, "MmWave vehicular beam selection with situational awareness using machine learning," *IEEE Access*, vol. 7, pp. 87 479–87 493, 2019.

[136] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *The IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015.

[137] A. F. Agarap, "Deep learning using rectified linear units (ReLU)," 2018, arXiv:1803.08375.

[138] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[139] C. M. Bishop, "Pattern recognition and machine learning," *Springer*, 2006.

[140] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Communications*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[141] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, "Hybrid architectures with few-bit ADC receivers: Achievable rates and energy-rate tradeoffs," *IEEE Transactions on Communications*, vol. 16, no. 4, pp. 2274–2287, Apr. 2017.

[142] 3GPP Technical Specification (TS) 22.261, "Service requirements for the 5g system," 3rd Generation Partnership Project (3GPP), Version 16.1.0, Oct. 2017.

[143] S. Lagen, K. Wanuga, H. Elkotby, S. Goyal, N. Patriciello, and L. Giupponi, "New radio physical layer abstraction for system-level simulations of 5G networks," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–7.

[144] O. Alliance, "O-RAN working group 2: AI/ML workflow description and requirements," Tech. Rep., Mar. 2019.