# Towards versatile access networks (Chapter 3)

Citation for the original published paper (version of record):

(article starts on next page)

# Towards Versatile Access Networks

*By Mir Ghoraishi, et al.*[1]

**now**
the essence of knowledge

---

1.    The full list of chapter authors is provided in the Contributing Authors section of the book.

## 3.1 Introduction

Compared to its previous generations, the 5<sup>th</sup> generation (5G) cellular network features an additional type of densification, i.e., a large number of active antennas per access point (AP) can be deployed. This technique is known as massive multiple-input multiple-output (mMIMO) [1]. Meanwhile, multiple-input multiple-output (MIMO) evolution, e.g., in channel state information (CSI) enhancement, and also on the study of a larger number of orthogonal demodulation reference signal (DMRS) ports for MU-MIMO, was one of the Release 18 of 3<sup>rd</sup> generation partnership project (3GPP Rel-18) work item [2]. This release (3GPP Rel-18) package approval, in the fourth quarter of 2021, marked the start of the 5G Advanced evolution in 3GPP [3]. The other items in 3GPP Rel-18 are to study and add functionality in the areas of network energy savings, coverage, mobility support, multicast broadcast services, and positioning [2].

The 6<sup>th</sup> generation (6G) cellular network will continue to modernize the existing radio access networks (RANs) towards satisfying 6G network's key performance

indicators (KPIs) [4]. Future networks will likely combine a range of RAN technologies from macro cells to small cells with very high-capacity, short-range links. This calls for dense small cell deployments, especially for throughput-demanding use cases required simultaneously by a high proportion of people in populated areas such as dense cities [5]. On the other hand, the technology's evolution towards cost reduction and improved efficiency requires all future deployment scenarios to rely on a superior transport network and network fabric that is flexible, scalable, and reliable to support demanding use cases and novel deployment options, such as a mixture of distributed RAN and centralized/cloud RAN enabled by artificial intelligence (AI)-powered programmability. The benefit of cell densification is to achieve a certain area capacity using less complicated hardware at the expense of using more APs (adding to the infrastructure requirement), which then means more interference. Distributed MIMO (D-MIMO) is a technology that combines the best aspects of ultra-dense cellular networks with MIMO technology to enjoy the strengths of both technologies [6].

From a technical perspective, an increased number of cooperating APs will pose new challenges for providing front/back-haul access to all nodes. On the other hand, at mmWave and sub-THz frequencies, a lot of bandwidth will be available (a total of 17.25 GHz of spectrum has been identified between 24 and 86 GHz [6, 8]), and links will be supported by very directive antennas limiting the amount of interference. This opens the door for integrated access and backhaul (IAB) solutions, sharing the same resources, but requiring potentially new beamforming and scheduling concepts [9]. It is foreseen that the wireless backhaul capacity will be enhanced in 6G by orders of magnitude, enabling the cost-efficient provision of 6G services using IAB even to remote areas that are currently cost prohibitive [5]. One cost-effective and backward-compatible method in densifying the cellular network is by combining different radio access technologies via multi-connectivity technology [10]. This technology was introduced originally as dual-connectivity in 3GPP Rel-11; nevertheless, it is still viable beyond the 5G era [11]. Another technology considered for 6G networks is reconfigurable intelligent surface (RIS), comprising an array of reflecting elements for reconfiguring the incident signals. Owing to their capability of proactively modifying the wireless communication environment, RISs have become a focal point of research in wireless communications to mitigate a wide range of challenges encountered in diverse wireless networks [12]. RIS technology is attractive due to its several advantages, namely, ease of deployment, spectral efficiency enhancement, and environmental affability [13].

There are yet several envisioned 6G use cases that require an extreme data rate in the order of 100 Gb/s even up to 1 Tb/s in specific scenarios. This cannot be fulfilled with current communication standards or realized by a simple evolution of available wireless technologies, rather it requires new technologies beyond the capabilities of

existing systems [14]. One approach for achieving such an extreme data rate relies on exploiting ultra-wide bandwidth of multiple GHz, which is available in the frequency range above 100 GHz, especially in the range 100–300 GHz, which is denoted as sub-THz [15].

The current chapter, "Towards Versatile Access Networks," provides an overview of the technologies introduced above, namely, D-MIMO, IAB, RIS, multi-connectivity, and sub-THz, in some more detail. The reader will notice that the chapter provides a non-exhaustive list of candidate technologies for the 6G RAN.

## 3.2    Distributed MIMO

### 3.2.1    What is D-MIMO for 6G?

Multi-antenna system technologies have evolved during every generation of wireless networks. After single-input single-output (SISO) systems, point-to-point MIMO has been introduced in the 3rd generation mobile communications (3G), and subsequently, the 4th generation (4G, or *long-term-evolution advanced*, LTE-Advanced) system is based on multi-user MIMO. Densification, i.e., increasing the number of base stations (BSs) per unit of space, is one of the main techniques that resulted in improving spectral efficiency in 4G and 5G cellular networks [1]. The drawback of this solution is high interference that negatively affects the performance of cell-edge users [15]. Networks of next generation will have to deal with even higher density of infrastructure to provide the expected performance [14]. This requires a rethinking of the underlying architecture to eliminate the cell boundaries [15, 17].

Heterogeneous networks (HetNets) have been utilized as an alternative to the homogenous cellular network architecture containing BS and possessing similar properties, where various transmission nodes (e.g., pico-cell, femto-cell, and remote-radio-head (RRH)) are installed within the same macro-cell area to improve the quality of service of cell-edge users as well as the system-wide service quality. These transmission nodes are network components that work in coordination with the central BS, but also differ from the BS in various aspects, such as transmit power levels and hardware. For example, RRHs are composed of antennas and RF amplifiers, and are connected to the BS via fibre cables or radio links using baseband signals, which transform the macro-cell deployment into a distributed antenna system concept.

The 5G NR standard introduces mMIMO, where each BS is equipped with many antenna elements, enabling it to serve numerous user equipment (UE) simultaneously by means of highly directional beamforming techniques. This approach is benefiting from channel hardening and favourable propagation utilizing the

deterministic channel and, hence, potentially eliminates the need for combating small-scale fading [18–21].

Joint Transmission Coordinated Multi-Point ( JT-CoMP), which enables coherent transmission from clusters of BSs to overcome the inter-cell interference within each cluster [20, 22], has also attracted the attention during the past 10 years; however, it did not become part of the 5G NR standard, as LTE standardization [23] did not deliver significant gains in practical deployments. This can be mainly attributed to the considerable amount of backhaul signalling for CSI and data sharing resulting from a network-centric approach to coherent transmission [24], whereby the BSs in a cluster cooperate to serve the UE in their joint coverage region. The practical implementation of JT-CoMP was also hindered by other attributes of LTE, such as frequency division duplex (FDD) operations and a rigid frame/slot structure, which did not allow for effective channel estimation.

One question for the 6G is what the next evolution of MIMO is going to be, and what the demands and needs of 6G networks from multi-antenna systems will be. 6G should provide limitless connectivity with both functional and deployment values [25]. Among other licenced bands, e.g., mid-bands, it will likely utilize mmWave frequencies and is expected to provide high spectral efficiency and predictable quality of service (QoS) to the UE. To ensure more consistent quality and non-intrusive, flexible, and robust networks, multi-point transmission is expected to become common [26]. It is envisioned that joint transmission and reception via spatially separated transceivers are going to be vital in upcoming systems [27].

Cell-free (CF) mMIMO [24, 28] combines the elements of small cells [20], mMIMO [21], and UE-centric JT-CoMP [20, 29], as illustrated in Figure 3.1. In a CF context, the mMIMO regime is achieved by spreading many antenna elements across the network (even in the form of single-antenna BSs [17, 30]), which provide enhanced coverage and reduced path loss. Moreover, a UE-centric coherent transmission extended to the whole network, where each UE is served jointly by several BSs, allows to practically eliminate the interference, as shown in [31].

Such a large-scale D-MIMO system, which can be thought of as the ultimate embodiment of concepts, such as network MIMO [31], multi-cell MIMO cooperative networks [22], virtual MIMO [32], ultra-dense networks, and JT-CoMP, is now regarded as a potential physical-layer paradigm shift for 6G networks [33].

A set of transmission nodes (APs) is assumed to be connected to a central processing unit (CPU) via fronthaul links, e.g., high-capacity fibre-optic cables, which convey both the UE-specific data and processing weights that enable network-wide processing for the computation of the AP-specific precoding strategies [24], see Figure 3.1. Since APs can perform channel estimation and distributed precoding locally, D-MIMO constitutes a scalable way to implement the network MIMO concept. Moreover, precoding can be designed by using channel estimates acquired
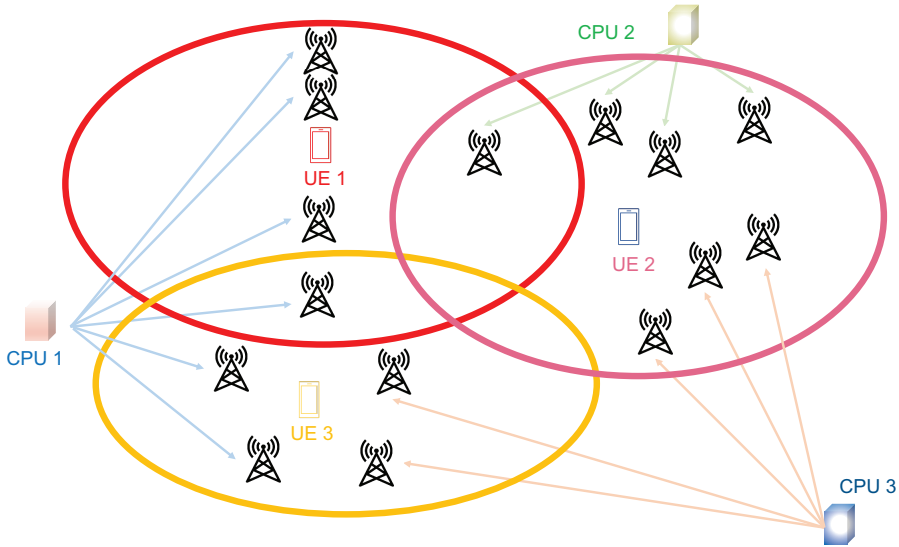
**Figure 3.1.** Illustration of distributed MIMO.

via uplink pilots by leveraging the channel reciprocity of time division duplex (TDD) operation; therefore, the overhead due to the channel estimation is independent of the number of APs and scales with the number of transmission layers. APs pertaining to different CPUs may serve a given UE.

As depicted in Figure 3.1, APs pertaining to different CPUs may serve a given UE, resulting in a distributed communication computation platform. Such a platform offers several features that are very well suited to support the diverse novel interactive use cases envisioned for 6G [34]:

1. Local connectivity-computational resources: many "mixed reality" applications rely to a significant extent on local content, and distributed processing can increase efficiency and reduce bottlenecks both regarding bandwidth and energy.
2. Proximity: excellent connectivity can be achieved with drastically lower transmit power. Applications can benefit from the distributed, local computational capabilities in the access infrastructure. The expenditure of overall network bandwidth can be drastically reduced as the dependence on medium- and long-distance connections will be minimized, resulting in very efficient usage of network bandwidth and energy. Interaction with energy-neutral devices essentially and realistically needs charging features to be close to these devices.
3. Redundancy: retransmissions can be avoided to achieve unperceivable latency and ultra-reliable connections.

4. Diversity: they can offer consistently excellent service levels that are essential for robust, innovative interactive applications, requiring imperceptible latency and zero outages. Precise and accurate indoor positioning can benefit from hyper-diversity, and favourable propagation conditions can be created to achieve consistent good communication quality and extensive spatial multiplexing.

The terminology for the transmission nodes in D-MIMO needs to be clarified, where both BS and AP are used in the literature; however, without expressing the functions placed in the units, any term like BS or AP does not provide sufficient information. Since BS can be interpreted with different functionalities, throughout this context, it is preferred to use the term AP as a distributed antenna or transmission node. A gNodeB (gNB) is introduced with functional splits in 5G era, e.g., radio unit (RU), distributed unit (DU), and centralized unit (CU), and the splits in D-MIMO have not been defined yet. One interpretation is that the APs may have little processing capability, and CPU may be the DU as it involves low-level functions. Accordingly, in such an interpretation, multiple CPUs may connect to CU. Herein, DU may be further split by centralizing the medium access control (MAC) and radio link control (RLC) layers with reference to the physical layer (PHY). In investigations, where the functional split is not fundamental, BS/AP and AP can be used.

## 3.2.2   D-MIMO Potential

At lower carrier frequencies (sub-6 GHz), where coherent transmission is possible, D-MIMO can be used to increase the spectral efficiency of the system and, in principle, to avoid inter-cell interference. Moving up in frequency, available bandwidth becomes larger, and spectral efficiency is not necessarily the main concern anymore. Instead, the reliability of the communication links becomes a primary concern. Reliability is impacted by the higher path loss, lower available output power of semiconductors, narrower antenna beams, and most importantly, a higher level of signal blockage. On the other hand, the feasibility of practical implementation highly depends on the RF hardware capabilities and other constraints, such as size, power source, and mobility. Moreover, the responses of different hardware components are influenced by the centre frequency, bandwidth, and waveform. Furthermore, the beamforming architecture and the possibility of exploiting spatial multiplexing depend on the radio channel characteristics, which need extensive measurement and modelling.

As, with D-MIMO, a link between the network and UE is provided by multiple APs, the likelihood that a link or combination of links with minimal blockage is available increases with *macro-diversity*. Hence, D-MIMO offers great potential to

exploit the spatial multiplexing gain of the channel and, at the same time, mitigate both unreliable links due to blockage and increased path loss. D-MIMO also allows for the densification of APs with, in the ideal case, no increase in interference. An increased density of operating APs will further reduce the likelihood of blockage and will also be necessary to have sufficient link margin due to output power limitations and increased path loss at mmWave and sub-THz frequencies.

### 3.2.3   D-MIMO: Roll-out Considerations

The appeal of D-MIMO as a 6G solution is its high degree of macro-diversity that results in a predictable service quality over the entire service area. The line-of-sight (LoS) probability for D-MIMO is very high, which makes it suitable for deployment in very high-frequency bands where radio propagation makes it challenging to provide a robust access link for mobile users. In theory, under the assumption of full CSI, there is no upper limit on the capacity of a D-MIMO system while densifying the distribution of nodes [31]. However, the capacity will be limited by practical constraints such as cost, power, and hardware impairments, as discussed in Section 3.2.4.

The main challenge for large-scale D-MIMO roll-outs is arguably the cost of installing many nodes in different places, each requiring fast and high-speed fronthaul connections. D-MIMO installations need to be easy to deploy, have a small and non-intrusive visual footprint, and be flexible to scale and extend. Clearly, there are seemingly conflicting requirements that D-MIMO needs to fulfil, as shown in Figure 3.2.

If every D-MIMO AP requires a dedicated fronthaul cable, then there is no way to make a large installation economically feasible. That is, in case every AP requires separate cables to each of its neighbouring APs, the installation would become complex, labour-intensive, and costly. Therefore, it must be possible to cascade multiple APs on the same fronthaul connection. Moreover, cascading or serializing the fronthaul may not be sufficient.
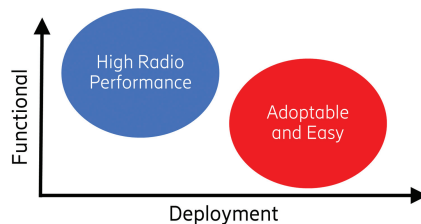


**Figure 3.2.** D-MIMO needs to fulfil seemingly conflicting requirements on functional complexity and deployment costs.
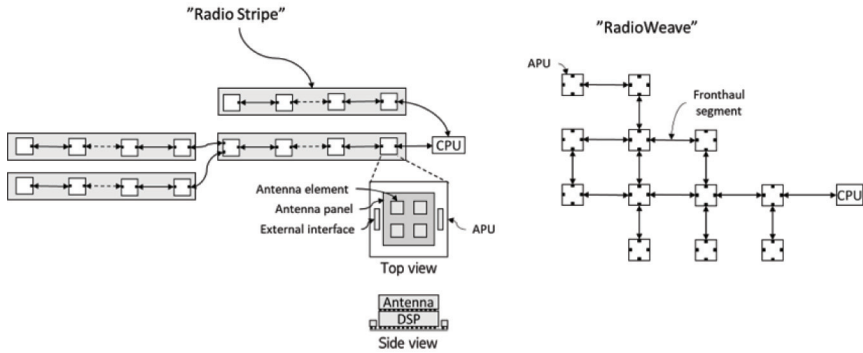
**Figure 3.3.** D-MIMO installations: (left) with radio stripe; (right) using both serialized (or cascaded) fronthaul connections plus parallel (or meshed) fronthaul connections. The fronthaul segments may be wired or wireless.

There are two different ways to provide fronthaul/backhaul to different APs, i.e., through fibre or wirelessly. For D-MIMO deployments, both ways need to be highly efficient. An optimized wired fronthaul can be realized through the so-called "radio stripes" [35], while wireless fronthaul/backhaul can be realized efficiently by integration with the access network. In this scenario, APs are integrated into a *radio stripe* as depicted in the left part of Figure 3.3. The radio stripe provides the fronthaul data and power supply for each AP as well as physical protection and encapsulation. With similar low-power techniques that are also used in mobile phones, the entire AP can be implemented in a system-on-chip (SoC) package that is denoted as antenna processing unit (APU) in Figure 3.3. An APU may contain some digital processing function (e.g., a digital signal processing (DSP) unit), an antenna panel consisting of one or more antenna elements, and one or more external interfaces that can connect to other APUs or to CPU. D-MIMO may also make use of a meshed fronthaul, sometimes denoted a RadioWeaves [36], as depicted in the right part of Figure 3.3. In a meshed fronthaul, the actual fronthaul segments may be implemented using wires, but the fronthaul segments may also be wireless. Further analysis of RadioWeaves is presented in Section 3.2.5.1. These architectures open possibilities to realize ultra-reliable links and interact with low-power devices. Moreover, capacity can be scaled up quite smoothly by adding more distributed components in the network.

Outdoor D-MIMO installations need to have a low visual footprint. With radio stripes, it is possible to hide the installation in existing construction elements in the environment, see Figure 3.4. To enable a small form factor of the APs or the APUs integrated inside a radio stripe, the power consumption needs to be very low. Fronthaul connections between APs or APUs can sometimes, but not always, be realized very easily with a cable. It is, however, not realistic to assume that every

**Figure 3.4.** Outdoor D-MIMO is well suited for dense urban areas, where it can provide invisible and high-capacity deployments.
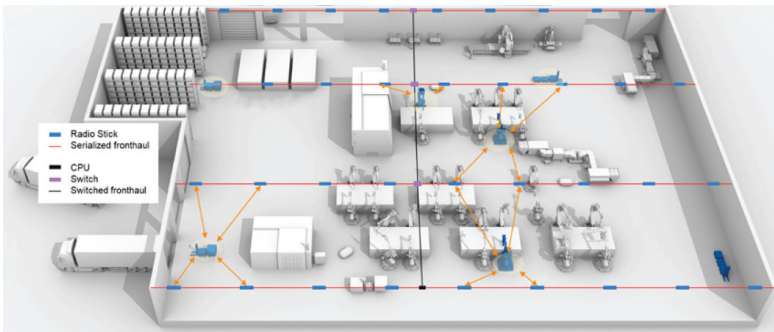


**Figure 3.5.** In indoor deployments, such as in a factory, D-MIMO promises to provide extremely reliable services with high capacity.

CPU can have a fibre backhaul connection. In dense urban areas, the backhaul connection can then be provided wirelessly by an existing macro BS.

Cascading or serializing the fronthaul connections is important for indoor installations as well, such as in a factory environment, as depicted in Figure 3.5. The concept can be compared to the electrical lights that are also installed in large volume inside a factory, where it is noted that this would not be possible if each electrical light required a separate cable installation. Just as one power cable can power up several lights, one fronthaul cable can connect several APs or APUs in a D-MIMO installation.

### 3.2.4   D-MIMO Deployment Considerations

#### 3.2.4.1   Hardware constraints

As it will be outlined in Section 3.6, signals transmitted at mmWave and sub-THz frequencies are subject to distortion by the transceiver hardware. This is even a more important issue in D-MIMO scenarios, as the received (or transmitted) signal

is collected (produced) from several hardware in separate locations experiencing different environmental conditions that may result in different effects on part of the received signal. In this section, the main hardware effects that are particularly relevant for D-MIMO systems are discussed, that is:

- Phase and frequency coherency
- Output power
- Reciprocity

To utilize distributed antennas to their maximum extent, coherent transmission and reception would be desirable. This requires all antennas participating in coherent transmission to be tightly synchronized. Going to sub-THz frequencies, coherent schemes will be very challenging to realize due to phase noise and frequency errors. Robust techniques against coherency errors must be developed, e.g., using antenna selection.

Another key hardware effect of current semiconductor technology at mmWave and sub-THz frequencies is that it is challenging to produce sufficient output power to maintain a robust communication link over a larger distance. As a result of increased power, more non-linear distortion is created. D-MIMO is a possibility to address this limited output power challenge: APs will have challenges to reach end-points due to output power limitations, but the inherent macro-diversity in D-MIMO and support for dense deployments enable more APs close to end-points. The effect of lower output power will result in transceiver designs with more parallel power amplifiers (PAs) combined with a higher number of antenna elements, yielding higher equivalent isotropic radiation power (EIRP). At lower mmWave frequencies where the EIRP, rather than AP output power, might be the limiting factor at certain deployment scenarios, D-MIMO has the capability to allow for lowering the EIRP by dense AP deployment and still to provide sufficient coverage.

To perform beamforming in a D-MIMO system, the channel information has to be received and known. One way to do this is to assume channel reciprocity, which means uplink and downlink channel are identical (in TDD mode). The channels estimated by a device contain the air-channel and also the analogue transceiver chains. Reduced coherence time (of channel and hardware) and components that are driven more non-linearly will make it more challenging to build D-MIMO systems based on channel reciprocity.

### 3.2.4.2   Architectural considerations

Among the desirable features, techniques for enabling scalable D-MIMO systems with converged access-backhaul-and-fronthaul, delivering extreme spectral efficiency, reliable access and robust mobility in scenarios ranging from low-bands

to sub-THz bands, taking hardware impairments and deployment options into account can be mentioned.

As a first step, it would be necessary to understand how much distribution is required and where is the sweet spot in terms of complexity versus robustness and performance considering the trade-off between distributed and centralized processing. Then, one needs to deal with the practical approaches to non-coherent operation in higher bands and transport solutions satisfying the requirements. Optimum solution would be phase-coherent transmission and one centralized processor, but it will be difficult to build and meet the feasibility requirements. The other extreme would be phase non-coherent transmission with duplicating every data in each AP and relying on single-frequency network, but it will be inefficient. Further research on finding the balance, in terms of complexity versus robustness and performance, is necessary. In addition, problems such as beam management aspects, practical approaches to non-coherent operation in higher bands, and transport solutions, e.g., wired/wireless, optical/electrical, and analogue/digital, satisfying the requirements need to be addressed.

At low-frequency bands (e.g., sub-6 GHz), the spectrum is scarce, thus there is a need to grow the spectral efficiency to increase capacity for the available limited bandwidth (e.g., 10–100 MHz). One way is to have a distributed coherent antenna system. In theory, D-MIMO and coherent multipoint transmission can enable higher capacity everywhere by adding additional coordinated APs. For these bands, the digital processing is feasible in the APs and can be carried out locally. It is only sort of impairments and practical limitations that will eventually limit the performance. Furthermore, distributed processing means that there will be less information, e.g., channel estimation and precoding, exchange between the APs and CPU. Since baseband data will be transferred, there will be less load (compared to RF modulated data) on the fronthaul that can be handled without high-capacity fronthaul links, e.g., a 10 Gb/s Ethernet digital fronthaul may be enough.

At higher frequencies, i.e., mmWave and sub-THz, relatively large system bandwidth, e.g., several GHz, are available, and there is not much traffic in early phases of the utilization of these bands, even the UE is limited in how much bandwidth it can process. That means there will be available spectrum that can be utilized for the fronthauling, and non-coherent transmission can be enough for early phases. Since APs are generally not nomadic, robust fronthaul links will be easy to maintain. Moreover, as frequency increases, antenna elements shrink, the antenna array will be relatively small, and beams become narrower. The main challenge in these bands is to realize and maintain a robust access link that supports mobility where the propagation environment is more challenging. Macro-diversity brings an advantage to achieve the high reliability, especially in NLoS, and a D-MIMO can benefit from a much higher degree of macro-diversity to overcome radio blocking in

case of narrow beams and weak signal penetration. Random blockage due to moving objects can be handled by redundant links. Nevertheless, digital processing, analogue-to-digital converters (ADCs) and digital-to-analogue converters (DACs) in these bands are power consuming and increase the chip cost and size of the processing unit. Small APs are essential in terms of deployment values, e.g., invisible, and aesthetic impacts, which is why it is required to move the digital processing from APs to the CPU, which may, as an option, mean that fronthaul will become analogue and needs to have a capacity on the level of fibre connection.

### 3.2.4.3  D-MIMO support in the ORAN

The D-MIMO network architecture intends to disaggregate the traditional CPU in multiple DUs, in line with 3GPP's 5G architecture [37] and to propose novel solutions based on fully distributed (aligned with ORAN Alliance specifications [38]), data-driven processing and local coordination. The disaggregation is vital for creating scalable versions of D-MIMO architectures [39], which will unlock the potential of deploying D-MIMO networking in future 6G networks with massive RU deployments.

In addition, in case of user-centric approaches, e.g., CF architecture, an important issue is cluster formation (selection of serving RUs). In contrast with the available simplistic distance-based solutions [29], it is important to dynamically allocate a sub-set (or cluster) of RUs to each UE based on (i) the radio propagation environment; (ii) quality of CSI estimates; (iii) constraints introduced by computation requirements; (iv) fronthaul links capacity; and (v) user mobility. Going a step further, innovative machine learning (ML) algorithms could be adopted for optimal cluster formation focusing on real-time operation by using historic data coming from the network, as well as for advanced modulation schemes and/or channel estimation/equalization. Finally, clustering RUs served by multiple DUs (in contrast to disjoint clusters in [39]) could provide an inter-DU coordination algorithm for decoding the actual signal. Moreover, exploring inter-DU coordination requirements and their effect in the spectral efficiency performance as well as dynamic adaptability of the coordination levels jointly addressing RU-DU and DU-DU coordination are some of the main research trends in the D-MIMO domain.

The D-MIMO architectures require fine granularity of the processing options (processing at AP versus CPU, inter-CPU coordination), which can be offered by hardware and software supporting the ORAN disaggregation concept [40]. In the following, the D-MIMO terminology is mapped onto the ORAN architecture, and a possible deployment is presented.

In ORAN architecture, the next-generation radio access network (NG-RAN) is disaggregated into ORAN CU (O-CU), ORAN DU (O-DU), and ORAN RU (O-RU). The O-CU is further split into O-CU control plane (O-CU-CP) and
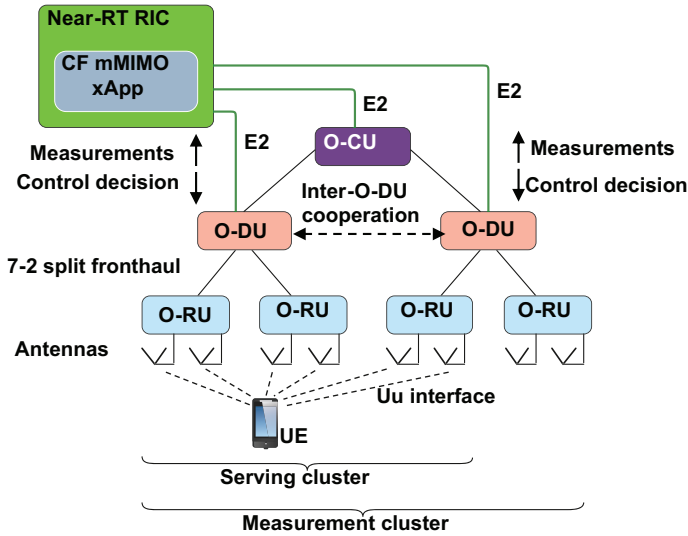
**Figure 3.6.** D-MIMO based on ORAN architecture, when a UE is served by O-RUs connected to several O-DUs, inter-O-DU cooperation may be enabled to boost the performance.

O-CU user plane (O-CU-UP). Multiple O-CUs and O-DUs are connected to the near-real-time RAN intelligent controller (near-RT RIC) for centralized NG-RAN performance control. Consequently, D-MIMO network components, particularly the CPU, can be associated with the ORAN nodes. The CPU is represented by the O-DU, while the AP is represented by O-RU.

The O-DU hosts the scheduler, which allocates radio resources for UE. Upon scheduler's decision, the user data are sent to the O-RU for transmission, or data are received at the O-RU. In the CF mMIMO, transmission/reception takes place simultaneously over multiple O-RUs, as shown in Figure 3.6. Thus, from the O-DU's point of view, UE can be divided into two categories as follows:

- **Local users**, who can be satisfyingly served by O-RUs connected to one O-DU.
- **Edge users**, who should be served collaboratively by O-RUs connected to two (or more) O-DUs. Serving of the edge users requires inter-DU cooperation (not yet in spec, but a studied interface [40]). It is also shown that serving edge users with O-RUs of multiple O-DUs improves the user spectral efficiency.

The main purpose of the inter-DU interface is to serve the user by having O-RUs connected to more than one O-DU, which is especially advantageous for edge users. One of the cooperating O-DUs should be considered a serving O-DU. The inter-O-DU cooperation interface can be used to both exchange user data between O-DUs as well as send necessary signalling required by the multi-antenna

processing (MAP). To avoid any confusion, the MAP terminology is used for practical deployment to describe the application of precoding and combining to the signals sent in downlink and received in uplink, respectively. Based on the recent studies, five D-MIMO deployment options with the support from ORAN architecture are identified [40]. The deployment options of the proposed approaches are discussed very briefly in Table 3.1.

Another challenge is how to apply the current state-of-the-art in user clustering for D-MIMO systems (e.g., in CF mMIMO literature) in an ORAN-oriented framework, which is done using Near-RT RIC. The ORAN specification defines the Near-RT RIC, which is responsible for optimizing the RAN performance. The RAN should work in the absence of Near-RT RIC. However, with the Near-RT RIC, the system performance improves, as it can be connected to multiple RAN nodes, where based on telemetry data from multiple E2 nodes, Near-RT RIC can optimize the RAN configuration to decrease interference and increase spectral efficiency.

The Near-RT RIC plays an important role in moving RAN away from cell-based design towards a D-MIMO architecture, in particular in a CF mMIMO scenario. In an mMIMO architecture, the number of antennas in an O-RU is high, but the number of O-RUs connected to O-DU is limited. In the CF approach, the number of antennas in an O-RU should be limited, but the number of RUs connected to an O-DU should be large. Each UE will be served by more than one O-RU, whereas O-RUs serving the UE can be connected to different O-DUs based on the location of the UE.

The Near-RT RIC based on E2 link capacity is connected to multiple E2 nodes, as shown in Figure 3.6. In the CF case, multiple O-DUs are connected to a Near-RT RIC and export the telemetry data of UE measured from multiple O-RUs to Near-RT RIC via E2 connection. As Near-RT RIC receives data from multiple O-RUs that can be connected to different O-DUs, it can cluster the O-RUs which should transmit to a particular UE. Two clustering approaches are considered for the CF approach:

- **Network-centric clustering (NCC)**, which consists in deploying fixed disjoint clusters of APs where the APs in a cluster serve only the UE residing in their joint coverage area.
- **User-centric clustering (UCC)**, which consists in deploying dynamic (possibly partially overlapped) clusters of APs based on the requirements of each connected UE in the system.

By analysing the performances, it is understood that UCC is superior compared to the NCC. In UCC, O-RUs serving the UE are grouped together based on signal strength of pilots received by the RU. Each O-RU measures the signal strength

**Table 3.1.** Deployment options.

| Deployment Option | Description of the Option | Inter-DU Coordination | ORAN Support |
|---|---|---|---|
| **Option 1** | The MAP vector is computed separately for each O-RU neglecting the channel between the user and antennas of other O-RUs | Absent (all O-RUs serve only the users located within the same O-DU) | Yes (Option 1 deployment is supported by current ORAN architecture) |
| **Option 2** | Same as Option 1, only difference here is that the edge users are served through the inter-DU coordination process | Present (all O-RUs of the O-DU serve local users. Additionally, the O-RUs serve edge users of neighbouring O-DUs) | No (Option 2 deployment is not supported by ORAN architecture) |
| **Option 3** | The MAP vector used by each O-RU for a particular user is computed jointly for all O-RUs of one particular O-DU (O-DU level information) | Absent (all O-RUs serve only the users located within the same O-DU) | Yes (Option 3 deployment is supported by current ORAN architecture) |
| **Option 4** | Same as Option 3, but only difference here is with inter-DU coordination interface present in the architecture: (i) all O-RUs of the O-DU serve local users; additionally (ii) the O-RUs serve edge users of the neighbouring O-DUs | Present (all O-RUs of the O-DU serve local users. Additionally, the O-RUs serve edge users of neighbouring O-DUs) | No (Option 2 deployment is not supported by current ORAN architecture) |
| **Option 5** | Option 5 may be implemented by connecting all O-RUs to one global O-DU where the global processing is performed | No (inter-DU coordination is not required where all the UE serve by a single DU) | Yes (ORAN can support this, but it is impractical due to high requirements on the O-DU hardware resources (processing, storage, and network) |

of pilot for a particular UE and sends it to O-DU, which in turn forwards it to the Near-RT RIC over E2 interface. The Near-RT RIC collects all the telemetry data from multiple O-RUs that can be from more than one DU. If inter-DU communication is supported, RU clusters can be formed independently of O-RU to O-DU connection for a particular UE. The cluster information then can be transmitted to the primary O-DU of the UE. This primary O-DU is then responsible for transmitting and receiving UE data from the clustered O-RUs.

The Near-RT RIC uses signal-to-interference-plus-noise ratio (SINR) of the pilot from a particular O-RU, O-RU load, and O-DU load to cluster O-RUs for a particular UE. Location service can also be used to determine the location of the UE in the network and be considered as a factor for clustering of O-RUs. As pilot measurement is periodic, the clustering is also done in a periodic manner. ML can be used to predict the UE movement based on its current SINR and previous SINR values and, in general, change in the SINR over time. This can help in predicting the UE movement towards a particular direction which, in turn, can help to cluster O-RUs serving the UE.

### 3.2.4.4   Analogue centralized beamforming

The type of D-MIMO system architecture and functional split that can be used depends very much on the overall system bandwidth. Therefore, on lower bands, where the bandwidth is limited, it can be feasible to use an electrical-based fronthaul technology such as Ethernet. Ethernet provides up to 10 Gb/s rate and can also provide up to 100 W or electrical power using PoE 802.3 bt type 4 [41]. On lower frequency bands, the digital processing requirements of each APU are also reasonably limited.

However, on sub-THz frequency bands, the very wide system bandwidth prohibits the use of electrical cables for the fronthaul, and we must instead use optical fronthaul connections. It is also not feasible, at least within the early 6G timeframe, to assume that the APUs will be capable of performing any meaningful digital processing on such high bandwidth signals. The APUs need to use low power to remain small enough for large volume installation, and this is not compatible with extreme requirements on digital processing.

Therefore, one promising architecture for high-frequency and high-bandwidth D-MIMO systems is to utilize analogue-radio-over-fibre and is depicted in Figure 3.7. To keep the APUs operating with low power usage (to reduce the heat dissipation and the resulting size and weight), the APU consists of an optical-add-and-drop-multiplexer, a photo-detector, a directly-modulated-laser, and some analogue components related to transmitter and receiver. Using wavelength-division multiplexing (WDM), it is still possible to serialize the fronthaul and connect many APUs to the same physical fibre cable. All digital processing (e.g., precoding and channel
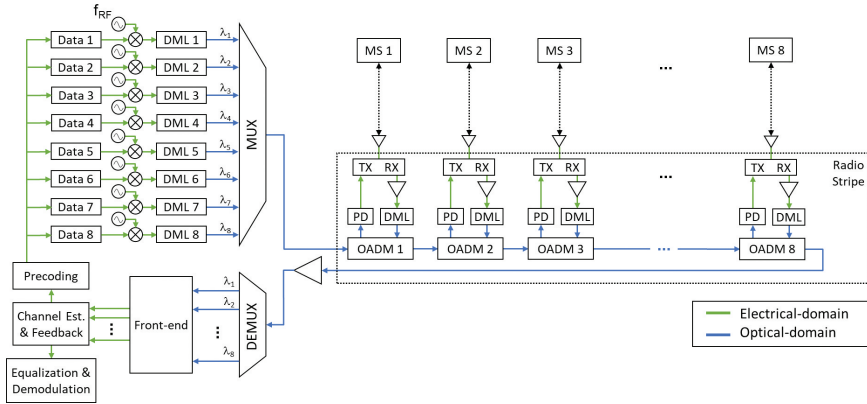
**Figure 3.7.** A wide bandwidth D-MIMO system can be implemented using analogue-radio-over-fibre and WDM.

estimation) is performed in the CPU where the constraints on size and weight are more relaxed, and a somewhat higher power consumption can be acceptable.

### 3.2.5  Some Recent Analysis of D-MIMO Scenarios

This section presents some recent results as a first step towards understanding how to develop a common scalable and frequency-agile architecture for practical D-MIMO deployments. In the following subsections, the theoretical analysis of a RadioWeaves deployment followed by the performance comparison of a distributed versus centralized zero-forcing as well as an analysis of a hybrid precoding is discussed.

#### 3.2.5.1  RadioWeaves deployment analysis

A preliminary study compared the connectivity in a RadioWeaves, as illustrated in Figure 3.3, infrastructure with a deployment based on one central "candelabrum" array with an equal number of antennas [42, 43]. The two deployment scenarios are shown in Figure 3.8. This analysis has considered 200 simultaneous, randomly located, users, and different topologies for the RadioWeaves infrastructure. It has assessed the transmit power requirement needed to guarantee that all users simultaneously receive 4 Mbit/s (200 × 4 b/s/Hz spectral efficiency).

The results are reported in [42]. The randomness in the distribution stems entirely from the randomness of the user locations. A significant gain can be achieved, that is, with a total transmit power of −20 dBm, the distributed RadioWeaves deployment on four walls will decrease the probability of not achieving the targeted throughput on a random location with a factor of 1000 compared to the candelabrum-based infrastructure operating with 0 dBm total output power.
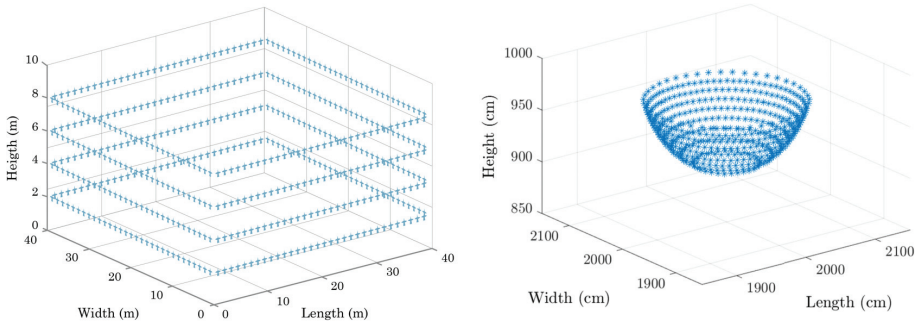
**Figure 3.8.** Distributed RadioWeaves infrastructure and central "candelabrum" deployment scenarios [42].

The RadioWeaves deployment achieves this 1000 times increase in QoS with 100 times less transmit power. This is shown in particular with respect to a central array in a candelabrum topology, which in itself is a quite good case and does create favourable conditions in a deployment with a central array.

The aim of RadioWeaves technology is to realize its great potential both in terms of performance metrics and energy efficiency [43]. Algorithm-architecture co-design is pursued, with specific attention for the bottlenecks in interconnects and mixed fronthaul-backhaul requirements [34].

### 3.2.5.2   Beamforming computations and weight distribution

When enough bandwidth is available at the fronthaul, a centralized digital beamforming can be applied. The fully digital beamforming is advantageous (compared to the analogue beamforming) since all the signal processing is done digitally. This can facilitate a more flexible design by using more degrees of freedom compared to the analogue beamforming introduced in Section 3.2.4 [44]. This section presents reciprocity-based (UE-centric) interference-aware distributed zero-forcing (IADZF) method and compares with centralized zero-forcing (CZF). Each distributed precoding done in clusters/subsets is interference-aware, i.e., it considers minimizing the power interfering to other clusters.

Let $N$ APs coherently serve $K$ randomly distributed UE in a UE-centric way. For UE-centric AP clustering ("subset"), the method in [41] has been set by grouping the APs that provide sufficiently high signal-to-noise ratio (SNR) values with the best channel quality that contribute at least $\alpha$%, e.g., 95%, of large-scale fading coefficients towards the $k$-th UE.

For the CZF, all APs form a single cluster. This means APs can transmit signals to any UE, and ZF is done in the central entity. In the distributed zero-forcing (DZF) case, multiple clusters are formed in a UE-centric way where APs in each cluster

**Table 3.2.** Simulation parameters.

| Parameter | Model Specification | |
|---|---|---|
| Frequency range (GHz) | 28 | 100 |
| Bandwidth (MHz) | 200 | 5500 |
| Maximum AP transmit power (dBm) | 13 | |
| Number of subarrays | 1 | |
| Vertical/horizontal antenna elements | 4/8 | 8/16 |
| Number of UE | {20, 40} | |
| Propagation Model | 3GPP InH [45] | |
| Area size | 100 m × 100 m | |
| Number of blockers | 1000 | |
| Blocker size | Max: 2 m × 3 m, Min: 0.5 m × 1 m | |
| Duplexing | TDD with 50% downlink | |
| AP noise figure (dB) | 7 | 10 |
| Overhead ratio | 1:3 | |

only serve the UE inside the given cluster. The precoding coefficients are calculated for each cluster separately. Performance has been evaluated in an indoor scenario with randomly distributed blockers, for different number of regularly deployed APs each with $M$ antenna elements and $K$ uniformly distributed single-antenna UE. The configuration parameters used in the simulations are given in Table 3.2. Perfect channel estimation and lossless, high capacity fronthaul is assumed. RF imperfections, hardware impairments, phase noise, and PA non-linearities are omitted from the scope of this work.

Figure 3.9 demonstrates that IADZF method applied over UE-centric clusters with $\alpha = 95\%$ clustering method can achieve the performance of CZF method, at which all APs serve each UE for different deployment sizes on sub-THz band. Two methods converge in relatively dense deployments. Due to the higher path-loss and weaker signal penetration, SE decreases as frequency increases.

IADZF has been compared with interference unaware TDZF in Figure 3.10 for different cluster sizes, $N_k = 1$, $N_k = 4$, $\alpha = 95\%$ clustering approach, and $N_k = N$, i.e., all AP simultaneously transmit to all UE, operating at 28 GHz. It has been shown that larger subsets using IADZF method bring more precoding gain and increase the spectral efficiency (SE). However, if distributed precoders are interference unaware, increased subset size degrades the performance.

Cumulative distribution functions (CDFs) of SINR values of 20 UE served by 49 APs for different serving cluster sizes, e.g., $N_k = 1$, $N_k = 4$, $\alpha = 95\%$ clustering approach, and $N_k = 49$ are shown in Figure 3.9. It has been shown that more
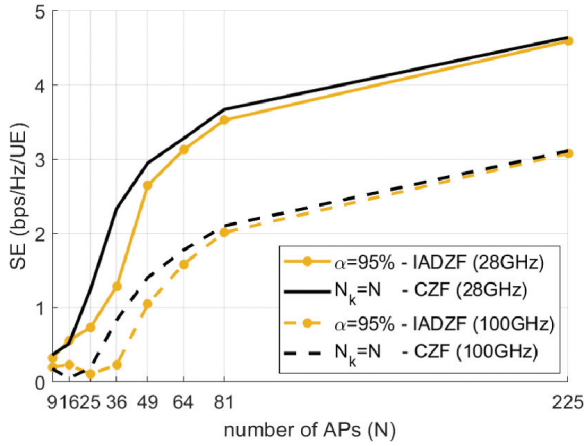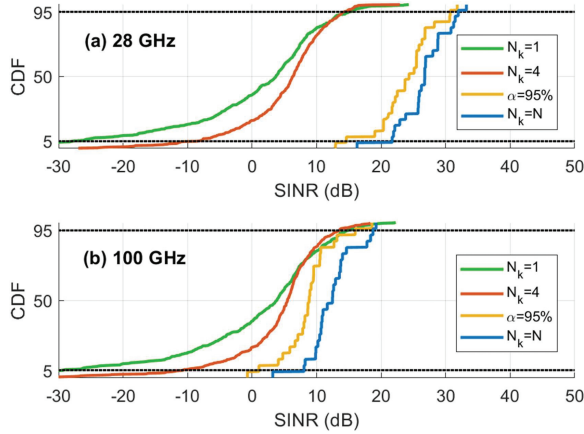
**Figure 3.9.** SE performance comparison of IADZF and CZF for different cluster sizes for 20 UE operating at 28 and 100 GHz.
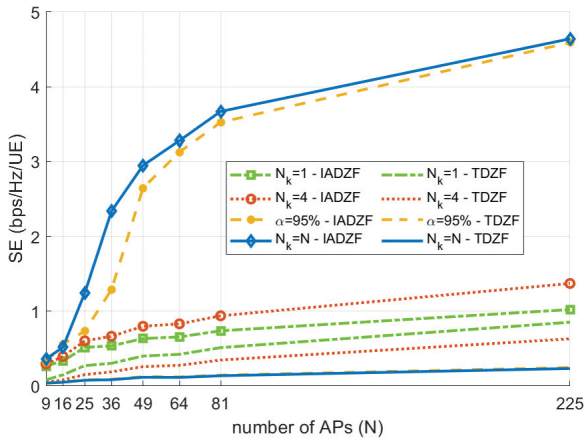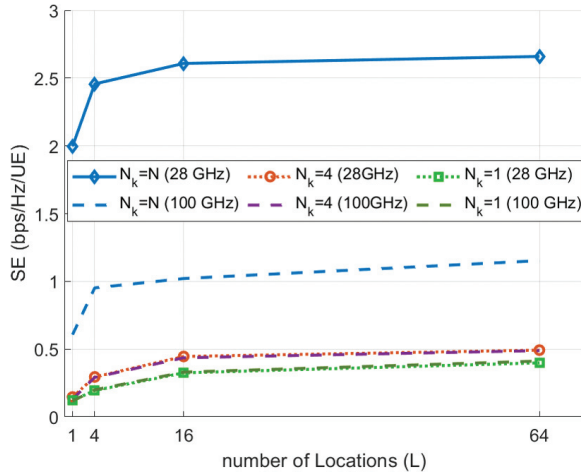


**Figure 3.10.** SE Performance comparison of IADZF and interference unaware TDZF for different cluster sizes for 20 UE operating at 28 GHz.

APs involved in coherent transmission bring higher degree of diversity, hence better interference cancellation and performance. Effect of the operating frequency is negligible for small size of serving AP subsets. Nonetheless, as frequency increases, APs close to the UE become dominant while distant APs do not contribute much, which decreases the received signal strength, eventually the SINR. Small clusters can still provide good enough SE performance on the average, and particularly can utilize the high bandwidth at high-frequency bands. Another highlight is that more APs can enable the cell-edge UE, e.g., 5% of the UE, having positive SINR.

Figure 3.12 demonstrates the average SE performance over 40 UE with reference to different antenna distributions, e.g., $N = 64 \times M$ antennas are distributed over

**Figure 3.11.** SINR performance for IADZF for $N = 49$ APs and $K = 20$ UE evaluated at (a) 28 GHz and (b) 100 GHz.



**Figure 3.12.** Performance comparison for collocated ($L = 1$), partially ($L = \{4, 16\}$), and fully distributed ($L = 64$) deployments for different operating frequencies for $N_k = \{1, 4, 64\}$.

$L = \{1, 4, 16, 64\}$ locations. It has been shown that semi-distributed ($L = 4$) or fully distributed ($L = 64$) deployments perform similar for single cluster case. Herein, there is a trade-off between implementation and deployment complexities. One cluster including all APs in the network can still achieve high SE values by collocating the antennas ($L = 1$), but it has higher implementation complexity. Many but smaller subsets ($N_k = \{1, 4\}$) are easier to implement joint processing; however, more scattered deployment is necessary for good enough performance.

### 3.2.5.3   Hybrid precoding in cooperative mmWave scenarios

The BS coordination schemes in mmWave networks that allow multiple streams to be transmitted jointly from multiple BSs are investigated by taking hardware and channel characteristics into account. Users are being served using spatial multi-flow and implements successive interference cancellation to decode the data streams sequentially.

In such a scenario, a downlink multi-cell and multi-user mmWave network with $M$ multi-antenna BSs and $K$ single antenna users are considered. It is assumed that hybrid precoding architecture is used at the BSs, to achieve high spectral efficiency with reduced hardware power consumption compared to that of the fully digital precoding, due to the reduced number of RF chains [46]. The number of antennas at BS is denoted by $N_m$ and the number of RF chains $L_m$. Depending on the number of phase shifters, the fully connected hybrid precoding architecture requires each RF chain connected to all antennas and the partially connected hybrid precoding architecture connects each RF chain to a subset of the antennas [47]. In addition, it is assumed that the sum power consumption of all BSs is divided into the hardware power consumption, including the phase shifters and the DACs, and the RF transmit power.

The objective is to minimize the sum power consumption such that the per-user minimum spectral efficiency, the per-BS maximum power constraint, and the hybrid precoding constraint are guaranteed. A sub-optimal algorithm by decoupling the optimization problem into an analogue precoding problem is proposed that only depends on the channel information, and a digital precoding problem that minimizes the sum power consumption by solving a semi-definite program. The proposed hybrid precoding algorithm jointly associates users to the BSs, finds the optimal BS silence strategy with minimum power, and enables us to jointly serve a user by multiple BSs. For the detailed system model and algorithm, cf. [48].

In Figure 3.13, the sum RF transmit power and the sum power consumption against the number of BSs are shown. In the scenario of this simulation, for each architecture, results based on all BSs being active are compared to the optimal case when the silent mode is enabled. Network parameters are as follows:

| Description | Parameter | Value |
|---|---|---|
| Number of antennas | $N_m$ | 64 |
| Number of users | $K$ | 4 |
| Number of RF chains per BS | $L_m$ | 4 |
| PS power consumption | $P_{PS}$ | 40 mW |
| DAC power consumption | $P_{DAC}$ | 200 mW |
| RF chain power consumption | $P_{RF}$ | 40 mW |
| Target spectral efficiency per user | $\tau_k$ | 4 bit/s/Hz |
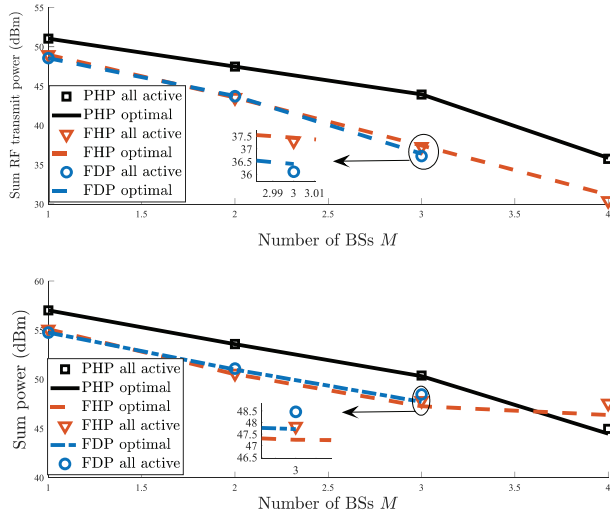
**Figure 3.13.** (a) Sum RF transmit power and (b) sum power consumption versus the number of BSs.

In Figure 3.13(a), the result shows that network densification and cooperative transmissions lead to a better chance for a user to be served by BSs with good channel conditions, thus, requiring less RF transmit power to achieve a target spectral efficiency. In Figure 3.13(b), for $M > 4$, the partially hybrid precoding starts to consume less sum power than the fully hybrid precoding, since the hardware power increase has less effect on the sum power consumption than that of the fully hybrid precoding. Also, Figure 3.13 shows that the power consumption difference between the optimal case and the sub-optimal case is small. It is worth noting that the total power consumption depends on the power consumption of the phase shifters, RF chains and the DACs, and the power scaling in the silence mode. Setting a low value for the hardware power consumption will give advantages to the all-active case and, thus, reducing the difference between the case of all-active and the case of the optimal silence mode. This parameters setting is based on the carrier frequency at 28 GHz. When higher frequencies are considered, the available transmit power will be further limited. Thus, the proportion of the hardware power will most likely increase, and the advantage of silence mode may be more prominent.

## 3.3   Integrated Access and Backhauling

The main aim in the IAB network architecture is to facilitate the dense deployment of the modern network architectures without the need for a fibre connection to each BS, by using the same spectral resources and infrastructures to serve both UE

in access as well as the BSs in backhaul [50]. IAB networks reduce deployment costs by replacing expensive wired backhaul to each cellular BS and do not cause massive degradation compared to all-wire network in realistic scenarios, as validated in [49]. Thus, it is expected that IAB architecture will be part of any advanced cellular RAN infrastructure lacking fibre connectivity to the RUs, but in particular for a D-MIMO architecture, IAB could be the efficient solution to the backhaul/fronthaul challenge. The D-MIMO infrastructure is expected to have many links connecting the CU and DU to multiple RUs covering the designated area. For such a case, IAB can provide a cost-effective connectivity.

Specifically, if bandwidth is not the primary constraint, as for example in early deployment and at sub-THz frequencies, IAB can be an efficient approach to maintain flexibility and efficient deployment. Moreover, IAB offers an advanced and flexible solution with multi-hop communications, dynamic resource multiplexing, and a plug-and-play design for low-complexity deployments. Additionally, the higher spatial reuse due to directivity in mmWave band and multi-beam systems and MIMO reduces cross-link interference between backhaul and access links allowing higher densification [49, 51].

Figure 3.14 shows a general IAB scenario, where the access and the backhaul of each BS need to share the same resources. BS1, master BS (MBS), is directly linked to the core network, CN, (e.g., by fibre backhaul), while BS2 and BS3 (secondary BS, SBS) use in-band backhauling [50]. In the network, there are constraints applied to all BSs and UE, notably, UE can see at least one BS. Besides, BS2 and BS3 can communicate with each other, while all BSs work in the TDD manner, i.e., either in transmit or receive mode in each time slot. In any case, one UE is served by only one BS in receive mode in each time slot, i.e., antenna/AP selection is used as the most primitive form of D-MIMO. One BS allows multiple access links in one time slot using different frequency resources. Note that this will increase the cost of link when scheduled and routed.
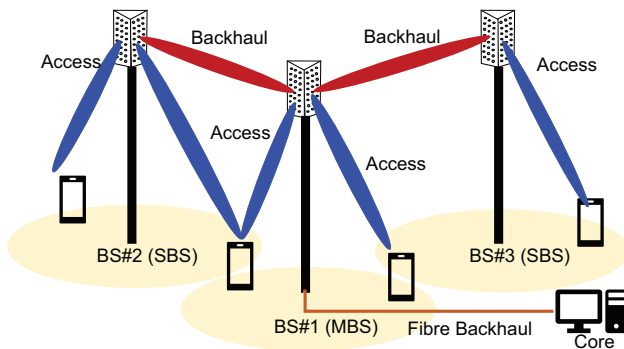


**Figure 3.14.** 5G Self backhaul-integrated access and the backhaul concept.

In this section, the efficient AP selection combined with access and backhaul scheduling in a TDD system where backhaul/fronthaul share resources is investigated.

### 3.3.1   IAB in 3GPP

To provide high data-rate requirements for backhauling a 6G BS, mmWave and sub-THz bands are expected to be used. In this case, backhaul links could be vulnerable to blockage, e.g., due to moving objects, seasonal changes (foliage), or infrastructure changes (new buildings). Thus, from a resilience perspective, it is important to ensure that an IAB node can continue to operate (e.g., provide coverage and end-user service continuity) even if an active backhaul path is degraded, lost, or even overloaded and failed. For this purpose, 3GPP has agreed on dynamic topology for IAB networks to autonomously reconfigure the backhaul network to achieve optimal backhaul performance under the above-mentioned circumstances [52], by supporting the autoconnection of an IAB node, network topology adaptation, and redundant connectivity. Moreover, the architecture is scalable, so that the number of backhaul hops is only limited by network performance (single and multi-hop backhauling are supported) [53].

In 5G NR IAB architecture, the network nodes are split into two types, namely: IAB-nodes and IAB-donors. Multiple IAB-nodes use wireless backhaul, and IAB-donors have fibre connectivity towards the CN. IAB-nodes and IAB-donors can serve UE and other IAB-nodes. Each IAB-node hosts two NR functions [49]:

1. Mobile termination (MT): for the wireless backhaul connection towards an upstream IAB-node or IAB-donor.
2. DU: for the access connection to the UE or the downstream MTs of other IAB-nodes.

Depending on using an AP or a BS (according to the definition provided in the previous section), a functional split of the radio protocol stack could be implemented with the control and upper layers in the IAB-donor CU, and the lower layers in the DUs of the IAB-nodes.

There is ongoing research regarding path selection techniques for 5G NR IAB networks. Different path selection techniques using a distributed approach are presented in [54], and the investigation of the performance in terms of hop count and the bottleneck SNR is analysed in [54, 55].

IAB adds new challenges in the task of scheduling to allocate resource between the access and the backhaul in IAB-node. Moreover, in a dense city where uplink/downlink traffics are constantly changing, scheduler with a fixed resource allocation could prevent the 5G network to realize its full potential.

Features of the scheduler task in the 5G IAB network is reviewed in [55], where the scheduling problem is split into two sub-problems and the ML task is defined for each sub-problem, namely, for the network access and the network backhaul. Finally, a smart scheduling solution based on deep learning is presented to address the problems above.

## 3.3.2   IAB Versus Fibre

In this section, the IAB network performance is discussed, and the evaluation of its performance in comparison with those achieved by hybrid IAB/fibre-connected networks using a finite homogeneous Poisson point process (FHPPP)-based stochastic geometry model, i.e., a Poisson point process (PPP) with a constant density, with random distributions of the IAB nodes as well as the UE inside a finite region is presented [56–58].

Figure 3.15 shows the service coverage probability of the IAB networks with those obtained by the scenarios having a fraction of fibre-connected SBSs, as
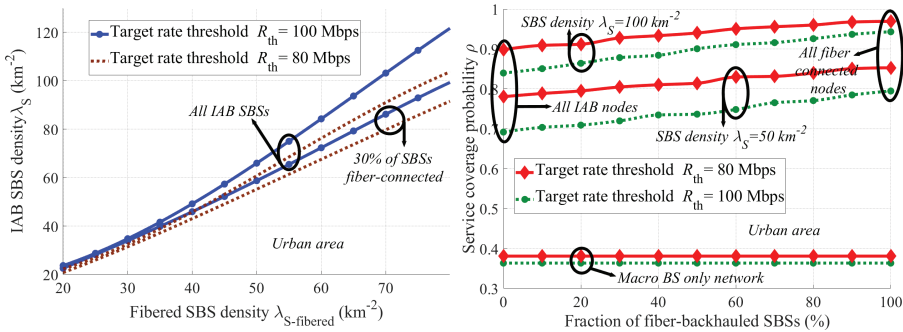


**Figure 3.15.** (left) Fibre-backhauled networks. (Right) Service coverage probability as a function of percentage of fibre-backhauled SBSs.

**Table 3.3.**  Simulation parameters.

| Parameter | Value |
| --- | --- |
| Carrier frequency | 28 GHz |
| Bandwidth | 1 GHz |
| Path loss exponents | [LoS, NLoS] = (2, 3) |
| Main lobe antenna gains | [MBS, SBS, UE] = (24, 24, 0) dBi |
| Side lobe antenna gains | [MBS, SBS, UE] = (−2, −2, 0) dBi |
| Antenna powers | [MBS, SBS, UE] = (40, 24, 0) dBm |

well as the cases without SBSs (refer to Table 3.3 for main simulation parameters). Figure 3.15-left compares the performance of IAB and fibre-connected networks, in terms of service coverage probability, i.e., the probability of the event that the minimum target rate requirements of the UE are satisfied. Moreover, Figure 3.15-right shows the network service coverage rate as a function of the fraction of fibre-connected SBSs and characterizes the system performance with the cases without SBSs.

This is motivated by the fact that some of the SBSs may have easy access to fibre. Here, it is assumed fibre-connected SBSs to be randomly distributed in the considered network area. It is observed that for a wide range of parameter settings, the IAB network can effectively provide the same levels of network service coverage probability as that of fibre-backhauled network with relatively small increase in the number of deployed IAB nodes.

Such a small increment in the number of IAB nodes leads to several advantages:

- **Increased network flexibility:** In contrast to fibre-backhauled networks, where the APs can be installed only in the places with fibre connection, the IAB nodes can be installed in different places if they have an acceptable connection with their parent nodes. This increases the network flexibility and the possibility for topology optimization remarkably.
- **Reduction in network cost:** An SBS is much lower in cost than laying fibre. Also, different evaluations reveal that, for dense urban/suburban areas, even in the presence of dark fibre, the IAB network deployment reduces the total cost of ownership.
- **Reduction of time-to-market:** Due to the required regulatory permissions, digging, and construction time, laying fibre typically takes a long time. In such cases, IAB can help to install new BSs/radio sites quickly.

### 3.3.3   Coordinated Mesh-based IAB

In a more general form, all BSs together form a backhaul/fronthaul mesh. Note that at this point, no assumption is made on the traffic to be routed to the core, i.e., the functional split between BS and core. Up to this point, this is abstracted in a cost-metric, which can be changed depending on beamforming capability, channel model, and deployment.

The overall problems to be solved are:

- Which AP should a UE connect to?
- In which time-slot should it be scheduled?
- Which route should the traffic be routed through fronthaul/backhaul mesh?
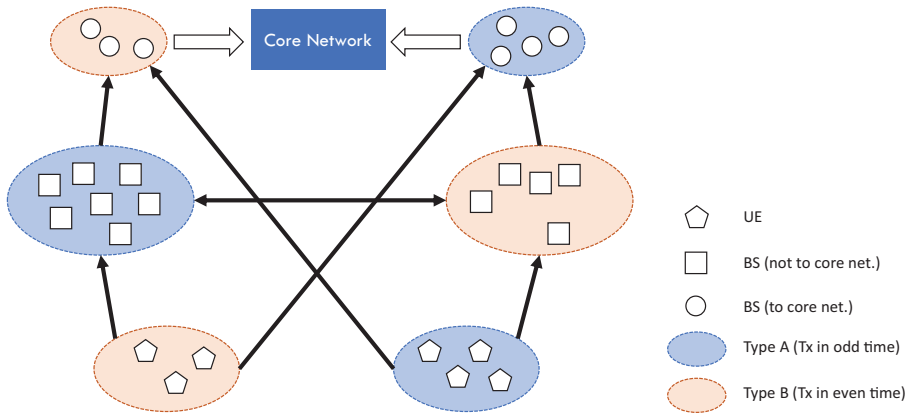
**Figure 3.16.** UE and BS/AP grouping in odd and even time slots, prior to routing optimization.

Given a cost metric (e.g., number of hops to core and throughput), the problem can be divided into time slot allocation and shortest-path optimization. Time slot allocation is illustrated in Figure 3.16, which shows how BSs and UE are grouped as part of optimization process.

For a given grouping of nodes (odd and even time slots), the second problem to be solved is between which nodes traffic should be sent. From which UE to which AP, and if an AP needs to transport traffic to the core, which backhaul/fronthaul links should be used?

Both problems need to be solved jointly, and exhaustive search becomes already prohibitively complex for small networks. Thus, a greedy algorithm was developed, which has a complexity that is linear in the number of UE.

### 3.3.3.1   IAB coverage analysis

As discussed in Section 3.2, the traditional cellular network will probably evolve to a kind of D-MIMO scenario, where it is likely to use APs with large distributed antenna arrays. The potential offered by IAB techniques can be very helpful in this regard by alleviating the fronthaul problem and by reducing the deployment cost in massive BS densification scenarios compared to the fibre deployment, which requires a noteworthy initial investment for installation.

Motivated by the presence of very wide bandwidths at mmWave carrier frequencies and above, IAB networks allow the operator to use part of the spectrum resources for wireless backhauling [53]. In 3GPP NR, IAB network configurations allow to provide flexible low-cost wireless backhaul using 3GPP NR technology in international mobile telecommunication (IMT) bands and provide not only backhaul, but also the cellular services in the same node. This will be a complement to existing microwave point-to-point backhauling in suburban and urban areas.

### 3.3.3.2   Genetic algorithm-based topology optimization for IAB

Due to the increase of network size in dense areas, which is the main point of interest in IAB networks, finding solutions for optimal network topology/routing is important. Since such optimization problem is very complex, an exhaustive search over all possible deployment options quickly becomes infeasible. This motivates a potentially suboptimal ML approach, which gives effective (sub)optimal solutions with reasonably satisfactory implementation complexity. Particularly, a genetic algorithm (GA)-based scheme to optimize the BS locations and non-IAB backhaul link placement is proposed. The details of the proposed scheme can be found in [59].

Unlike the non-IAB backhaul-connected networks, IAB networks may be prone to environmental effects, especially due to the blockage in dense urban environments and the tree foliage in sub-urban environments. It should also be noted that although IABs' main point of interest is dense urban areas, it has the capability to be deployed in suburban areas as well.

The evaluation of the blockage effect in urban areas and tree foliage on the network performance of an IAB network in suburban areas with random deployment and GA-optimized non-IAB backhaul link distribution is presented in Figure 3.17. Here, the results are presented for different rate requirements of the UE for the access links with path loss exponents 3 and 4 for LoS and NLoS propagations, and main lobe antenna gains of 18, 18, and 0 dBm for MBS, SBS, and UE, respectively. Particularly, Figure 3.17 (left) shows the service coverage probability considering the PPP-based germ-grain blocking model, while in Figure 3.17 (right) presents the results for the average hop distance of 450 m which corresponds to SBSs density of 8 $km^{-2}$ in a suburban area. In Figure 3.17 (left), the MBS, SBS, and UE transmit powers are as Pm, Ps, Pu = (40, 24, 0) dBm, and in Figure 3.17 (right) the tree density are $\lambda_M, \lambda_S, \lambda_U = (2, 50, 500)$ $km^{-2}$.
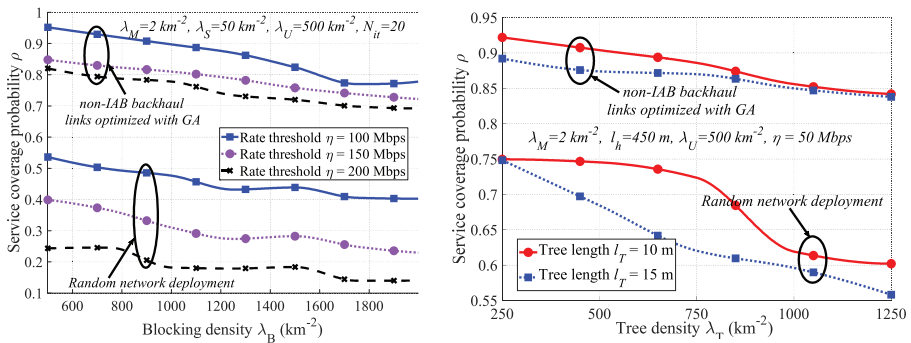


**Figure 3.17.** Service coverage probability of the IAB network; (left): as a function of the blocking density; (right) as a function of tree density.

Based on Figure 3.17, the following conclusions can be derived:

- The GA-based optimized network deployment shows considerable resilience to blockage and tree foliage, compared to the unoptimized random network.
- The service coverage probability in suburban area with a random IAB network deployment is considerably affected by the tree foliage loss. This is especially witnessed when the tree density is increased. However, with the introduction of GA-based optimization on selecting the appropriate non-IAB backhaul link distribution, a resilience to the tree foliage is observed.

In general, IAB robustness is hard to predict in the presence of tree foliage due to the influence of characteristics of the trees/vegetation on the link quality. Particularly, the link quality may vary due to the number of wet trees, snow on the trees, leaf percentage in different seasonal changes, and wind. However, it can be concluded that proper network planning can reduce the adverse effects on IAB, even though it is prone to medium/highly densified tree foliage in suburban areas. Moreover, mmWave IAB in areas with low/moderate amounts of tree foliage is expected to perform well.

## 3.4   Reconfigurable Intelligent Surfaces

Research in MIMO and D-MIMO context can be linked with another technology, i.e., RIS adoption. RIS adoption in telecommunications is a novel technology foreseen as one of the key enablers of the 6G mobile networks.

RIS is a two-dimensional surface of engineered material whose properties are reconfigurable rather than static, and with RIS, it is possible to shape how the surface interacts with wireless signals, enabling the wireless propagation environment to be fine-tuned [60]. The surface can include mainly passive elements without doing digital processing or any signal amplification. Nonetheless, some RIS surfaces can have relaying capability providing signal amplification. Moreover, RIS can have active elements that can enable digital signal generation, i.e., it can be interpreted as a sort of extra-large mMIMO.

RISs are low-cost and limited-power devices that can redirect, in a programmable fashion, the impinging waves towards the desired directions. Hence, providing control on the propagation environment in turn becomes an optimization variable to enhance communication link performance [13, 61]. RIS is suitable for a number of use cases, such as performance boosting, electromagnetic field exposure minimization, localization, and sensing [62]. Moreover, proposed RIS applications in the recent literature cover several application scenarios under

different assumptions. These scenarios can be categorized as [13]:

- RIS-enhanced cellular network, where RIS is used to establish the link between the users located in a blind spot and the BS. Thus, the QoS in Het-Nets and the latency performance in mobile edge computing (MEC) networks are improved [63] or can act as a signal reflection to support massive connectivity via interference mitigation in device-to-device (D2D) communication networks, or to strengthen the received signal power of cell-edge users and mitigating the interference from neighbour cells [64].
- RIS-enhanced unmanned systems, where RIS can be leveraged for enhancing the performance of unmanned aerial vehicle (UAV)-enabled wireless networks [65], cellular-connected UAV networks [66], autonomous vehicular networks, autonomous underwater vehicle (AUV) networks, and intelligent robotic networks by fully reaping the aforementioned RIS benefits.
- RIS-enhanced Internet of Things (IoT), where RIS is exploited for assisting intelligent wireless sensor networks, e.g., in intelligent agriculture and intelligent factory scenarios [67, 68].

In this section, first a proposed architecture based on introducing logical components for the control and orchestration of the RIS operation is discussed. Then, an analysis on the influence of the RIS position and orientation on its overall performance is presented, followed by discussions on two specific RIS scenarios, i.e., cascaded multi-RIS scenario and RIS-assisted UAV system.

### 3.4.1   Proposed Architecture for Efficient RIS Deployment

RIS can be organized for adaptive network configuration and orchestration depending on scenarios and application needs [13, 61]. The integration of RIS in the RAN foresees additional logical components, interacting with each other and with the legacy components, to enable integration and orchestration of such devices. A proposed architecture and corresponding components for the RIS-aided control in the RAN are depicted in Figure 3.18, where the following components are represented [68]:

- **RIS:** the RIS itself is the reflecting intelligent surface based on technologies such as reflect-array or meta-material. RIS devices are typically constituted by a grid of discrete unit cells spaced at sub-wavelength distance. The electromagnetic (EM) response of each unit cell can be controlled in a programmable manner by altering several EM response parameters, such as phase, amplitude, polarization, and frequency. Moreover, recently proposed RIS hardware designs are providing the surface with channel sensing capabilities [69, 70]. Each RIS deployed in the network is associated with a RIS

actuator that controls the RIS surface within a time granularity in the order of 10–100 ms.

- **RISA** (RIS actuator): the logical entity in charge of actuating the commands received from the RISC (RIS controller, refer to the definition below) and translate them in the configurations of the RIS, i.e., reflection properties of the surface. RISA can provide feedback to the RISC, e.g., in case of sensing capabilities at the RIS to provide the network with additional context information. The envisioned time granularity of the RISA is in the order of 1–20 ms.

- **RISC** (RIS controller): the controller associated to a RIS actuator that generates the logical commands associated with the switching between different states/configurations of the RIS elements (e.g., predefined or custom phase shifts configuration at the RIS). The RISC can embed third-party smart algorithms to derive the desired RIS configuration to be applied. Alternatively, the configuration can be setup based on received commands from other elements of the network. In the former case, the RISC takes care of the optimization of the RIS surface autonomously to form other network optimization operations, whereas in the latter case, the RISC behaves as an interface that controls the RIS based on external instructions, e.g., if the RIS configuration is optimized jointly with other network optimization operation. A RISC time granularity in the order of 20–100 ms is envisioned.

- **RISO** (RIS orchestrator): a logical component placed at a higher hierarchical level, to serve the role of orchestrating multiple RISCs in the network. Depending on the application, the time granularity of the RISO is expected to be in the order of 100–1000 ms.

As it is presented in Figure 3.18, there is a one-to-one correspondence between RIS and RISA. The RISA, RISC, and RISO can be virtualized, abstracted, and deployed into edge or central clouds, while physical devices (RF) shall be placed on site. The RISA operates directly on the RIS devices through the general interface called Open Environment that would accommodate heterogeneous RIS technology. Still depending on the RIS technology, different functions can be placed in the RISA to accommodate different RIS technologies, e.g., CSI feedback in the case of sensing capabilities at the RIS, etc. Therefore, the RISA provides an open interface to integrate heterogeneous RIS devices capabilities.

## 3.4.2  RIS Position and Orientation Influence on the Performance

As a challenge in the RIS-assisted scenarios, the quality of the communication is expected to depend strongly on the RIS placement in a RIS-aided link. This is
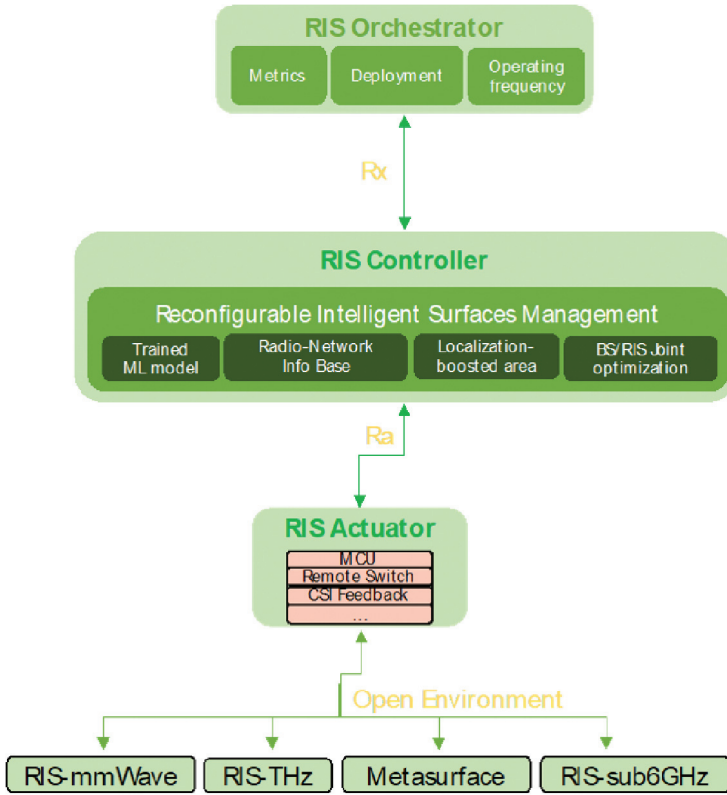
**Figure 3.18.** Proposed components and architecture for RIS control and orchestration.

because the path-loss changes with the relative distance of the RIS from the BS, the user, and their relative orientations. Therefore, the position and orientation of the RIS, with respect to the location of the BS and the user, need to be taken into account [71, 72].

Figure 3.19 shows an indoor scenario where the direct LoS link between the BS and the user is interrupted. The angular relation between the orientation of the RIS, the transmitted signal to the RIS, and the reflected signal is also clarified. The optimal RIS placement refers to the topology that guarantees a minimum threshold for the received power, for every possible location of a mobile user within the area of interest. To understand the impact of the RIS position and orientation on the received power, it would be instructive to investigate how the RIS captures the incident power and how it redistributes it in space. With respect to the incident beam, the power captured by the RIS depends on the illumination conditions, i.e., whether it is fully illuminated (therefore capturing only a part of the incident beam) or partially illuminated (capturing the entire incident beam). With respect to the reflected beam, assuming that the RIS is lossless, the total power of the beam
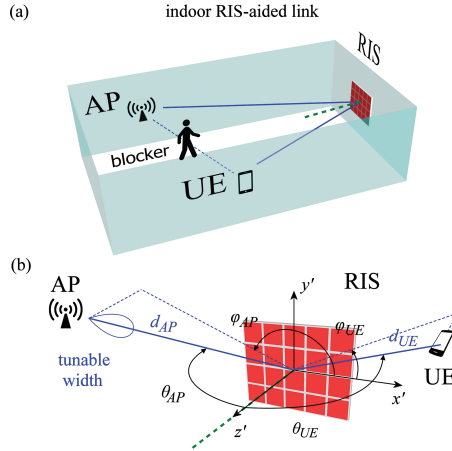
**Figure 3.19.** Indoor RIS-aided link system model: (a) LoS communication link interruption; (b) relative angles between the BS, RIS, and the user equipment.

reflected by the RIS is the same for all possible UE positions. However, the received power is determined by the local power density of the reflected beam (not the total beam power) and, therefore, beam spreading that causes the beam peak power to drop will have a direct impact on the received power.

A set of simulation results for the spatial distribution of received power as a function of user equipment position in ideal alignment is presented in Figure 3.20. Here, Figure 3.20(a) and (c) shows the received power as a function of the UE position in the absence of misalignment, while Figure 3.20(b) and (d) shows cross section at $z = 3$ m for the cases marked at Figure 3.20(a) and (c) with steering angles at $0°$, $20°$, and $40°$. Furthermore, the cases in Figure 3.20(a) and (b) represent the fully illuminated case, whereas Figure 3.20(c) and (d) shows the partially illuminated case. The contour lines represent the locus of user equipment positions, where the same received power can be achieved for each individual case. This is shown in Figure 3.20(a) and (c) for all possible user equipment positions within the illustrated area. The solid lines represent three chosen steering angles, namely steering angles at $0°$ (blue), $20°$ (orange), and $40°$ (yellow), and the dashed line marks their values at distance $z = 3$ m from the RIS (denoted with the short black line in the top). The contour lines represent the locus of user equipment positions at which the same received power can be achieved for each individual case. The spatial distribution of the received power for the chosen angles along the dash-line is shown in detail in Figure 3.20(b),(d). It is observed that, with increasing RIS to UE elevation angle, the beam spreading becomes more severe, resulting in the reduction of the beam's peak power and, consequently, of the received power (in case of ideal alignment the user equipment is located at the peak of each distribution).
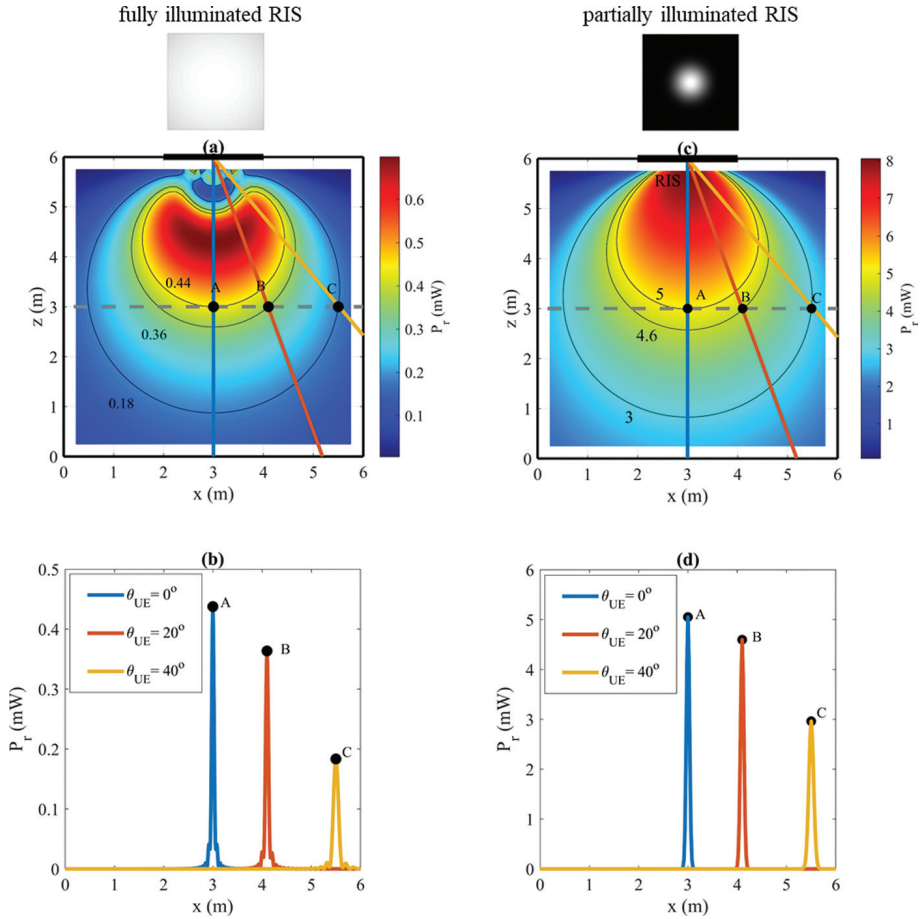
**Figure 3.20.** (a), (c) Spatial distribution of the received power as a function of the user equipment position in case of ideal alignment; (b), (d) cross sections at $z = 3$ m for the given steering angles; while (a) and (b) represent the fully illuminated case and (c) and (d) the partially illuminated case.

In the light of the above analysis, it is observed that the optimal RIS placement depends on both the RIS orientation and the illumination conditions and can be formulated in terms of the coverage provided by the RIS (or RIS efficiency).

As an example, the RIS orientation efficiency is shown in Figure 3.21 in terms of coverage and as a function of the orientation angle, i.e., the angle between the top wall and the RIS normal, as shown in the inset. While both BSs are equidistant from the RIS, the RIS orientation angle $\theta_o$ modifies the angle $\theta_{AP}$, with direct consequences on the room coverage. Figure 3.21(a) represents the fully illuminated case with BS antenna gain equal to 35 dB, whereas Figure 3.21(b) shows the partially illuminated case with AP antenna gain equal to 55 dB. For the fully illuminated case shown in Figure 3.21(a), the received power is proportional to $\cos(\theta_{AP})$. As a
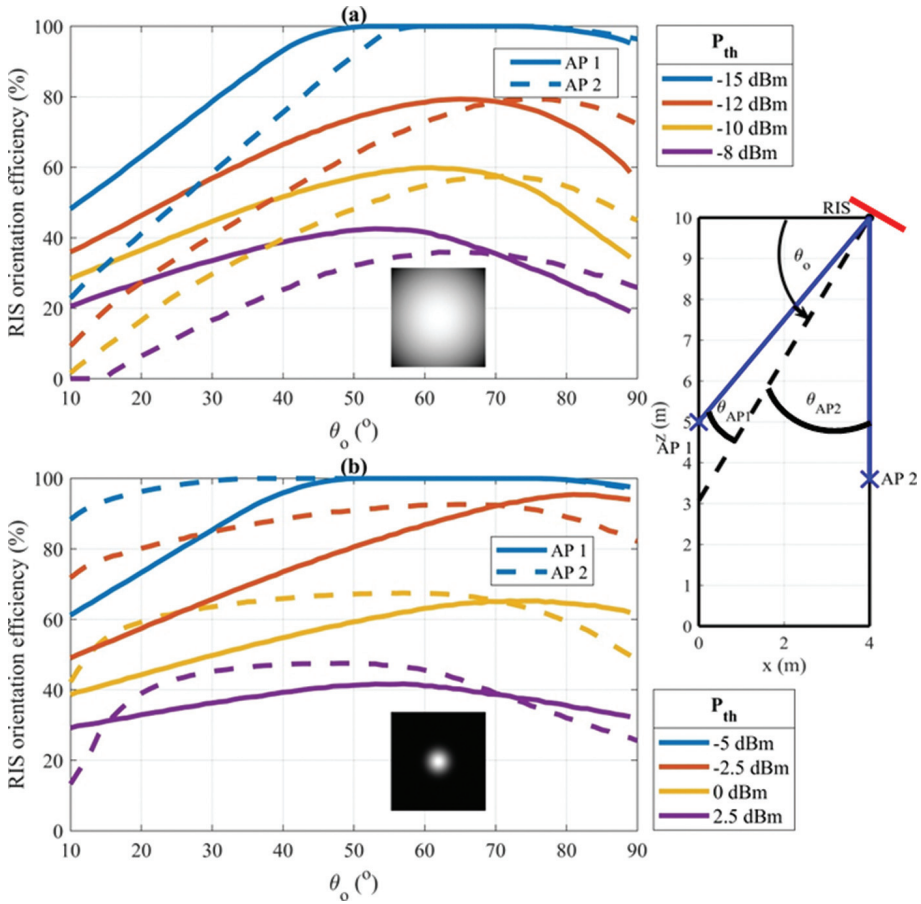
**Figure 3.21.** RIS orientation efficiency versus its orientation angle $\theta_o$. (a) Fully illuminated RIS (35 dB BS antenna gain) and (b) partially illuminated RIS (55 dB BS antenna gain), for the indoor scenario shown in (c).

result, the efficiency of the first BS is higher than the efficiency of the second BS when $\theta_{AP1} < \theta_{AP2}$ (i.e., for $\theta_o \to 10°$), and vice versa. The two efficiencies become equal when $\theta_{AP1} = \theta_{AP2}$, which occurs when the RIS normal points towards the opposite corner of the room (bottom left corner). On the other hand, for the partial illumination case shown in Figure 3.21(b), the efficiency increases with the increase of $\theta_{AP}$, because the beam footprint on the RIS becomes larger (acquires elliptical shape), in turn inducing weaker beam spreading and stronger peak power. In this case, the efficiency of AP1 is higher than the efficiency of AP2 when $\theta_{AP1} > \theta_{AP2}$ (i.e., for $\theta_o \to 90°$), and vice versa. Again, when $\theta_{AP1} = \theta_{AP2}$, the efficiency for both AP positions is equal. The solid and dashed lines illustrate the room coverage for the first and second BSs, respectively ($d_{AP} = 6.4$ m for both cases). The power thresholds examined in each case are depicted with the color-coded lines.

In conclusion, the RIS efficiency depends on its (a) position, (b) orientation, and (c) size relative to the size of the incident beam's footprint. In case of a fully illuminated RIS, the received power is proportional to cosine of the incident signal, which translates in decreased efficiency when the incident angle increases, and vice versa. The opposite occurs in the partially illuminated case, where increasing the incident angle leads to wider incident beam-footprint, in turn to weaker beam spreading, higher peak power and, consequently, to higher received power at the user equipment and therefore higher efficiency.

### 3.4.3   Cascaded Multi-RIS Scenarios

In case, the blind spot problem cannot be resolved using a single RIS (as, e.g., the LoS from the RIS to any of BS or the user is missing), then multiple-RISs can be used to provide uninterrupted connectivity between the BS and the user [73–75].

Multi-RIS can enable several novel functionalities, including, but not limited to, blockage avoidance, routing, coverage expansion, and beam splitting. An example scenario is a multi-RIS-empowered outdoor THz wireless systems discussed in [74] (Figure 3.22).

The result of a statistical characterization of the multi-RIS outdoor THz links and the associated performance gains, e.g., outage probability of the link as a function of RISs sizes and distances, are discussed in [73, 74]. These references analyse the impact of turbulence on the outage performance of multi-RIS-empowered THz
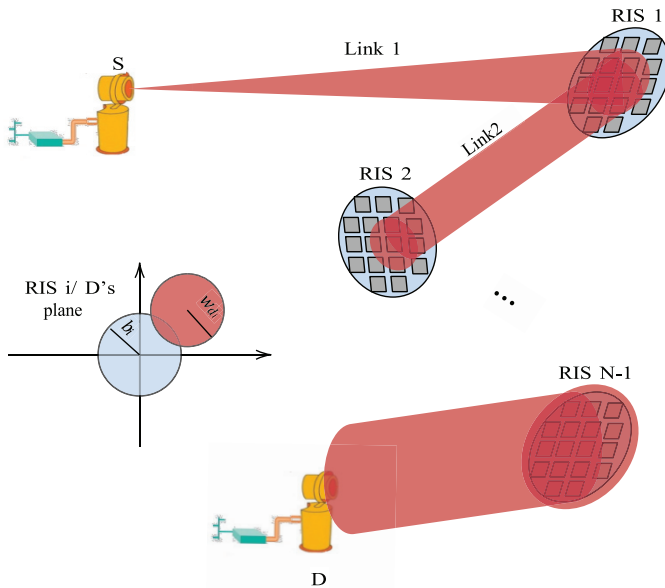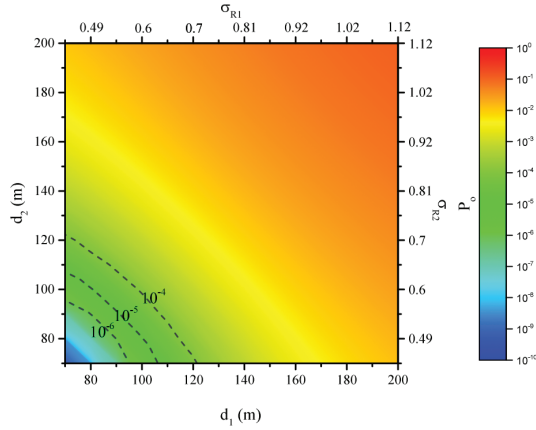


**Figure 3.22.** Cascaded multi-RIS-empowered THz wireless model.

**Figure 3.23.** Outage probability for a link consisted of two RIS at distances $d_1$ and $d_2$.

wireless systems in the absence of misalignment and hardware imperfections for a double RIS link setup. Figure 3.23 shows the outage probability of the link as a function of $d_1$ and $d_2$, which are the distances to the first and from the first to the second RIS. Moreover, $\sigma_{R1}$ and $\sigma_{R2}$ are defined as

$$\sigma_{R_i}^2 = 1.23 C_n^2 \left(\frac{2\pi}{\lambda}\right)^{7/6} d_i^{11/6}$$

with $\lambda$ is the wavelength and $C_n^2$ is the reflection index structure parameter [73, 74].

As expected, for a fixed $d_1$, as $d_2$ increases, $\sigma_{R2}$ increases; thus, an outage performance degradation is observed. Similarly, for a given $d_2$, as $d_1$ increases, $\sigma_{R1}$ also increases, i.e., turbulence intensity increases, in turn, the outage probability increases. Finally, from Figure 3.23, we observe that for a given transmission distance, $d_1 + d_2$, the worst outage performance is observed for $d_1 = d_2$.

## 3.4.4  RIS-assisted UAV Systems and Performance Analysis

As it was mentioned in the introduction of Section 3.4, one potential application of RIS is in enhancing the performance of UAV networks. Most existing analysis of the RIS-assisted UAV wireless systems, however, assumes that the RIS-UAV link is not directional. This results in neglecting the adverse effects of UAV disorientation and/or misalignment of the RIS-UAV beam. However, as the operating frequency and thus the directionality of links increase, even small disorientations and/or misalignments may adversely affect the performance of the RIS-assisted UAV wireless system. Aside from the joint effects of disorientation and misalignment, another important performance-limiting factor in high-frequency communications is the effect of transceiver hardware imperfections.
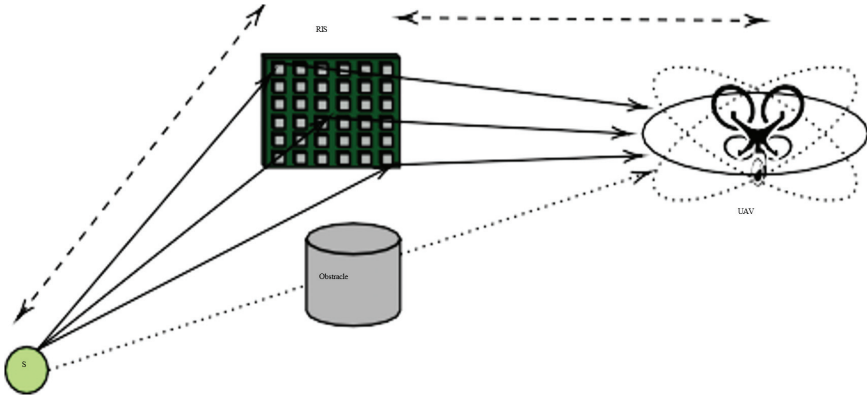
**Figure 3.24.** RIS assisted UAV scenario.

The performance analysis of RIS-assisted UAV wireless systems that accounts for the impact of different propagation environments, UAV disorientation and/or misalignment of the RIS-UAV beam, and the transceivers RF front-end imperfections is presented in [73]. Motivated by this, in this section, a contribution in this area is presented [76].

Figure 3.24 shows a RIS-assisted UAV scenario in which a BS (or a UE) communicates with a UAV via a RIS. For the sake of analysis, it is assumed that no direct link can be established between the BS and the UAV, due to blockage. It is assumed that the UAV is in a hover state, where both the position and orientation of the UAV are not completely fixed. In this case, both the BS and the UAV are equipped with single antennas, while the RIS is equipped with multiple antennas.

Figure 3.25 demonstrates the joint impact of hardware imperfections and fading on the outage performance of RIS-assisted UAV wireless systems, where the outage is plotted as a function of the transmitter and UAV receiver error vector magnitudes, $\kappa_s$ and $\kappa_d$, for different values of threshold [76]. In this case, $\kappa_s = 0$ indicates that the BS is equipped with ideal RF front-end. Similarly, $\kappa_d = 0$ means that the UAV is equipped with ideal RF front-end. Hence, the $(\kappa_s, \kappa_d) = (0,0)$ point represents the case in which both the S transmitter and UAV receiver are ideal. As expected, for given threshold and $\kappa_s$, as $\kappa_d$ increases, the outage also increases. Similarly, for fixed threshold and $\kappa_d$, as $\kappa_s$ increases the outage performance degrades.

## 3.5   Multi-Access Connectivity

The 5G CN supports integration of non-3GPP access networks via interworking functions that provide a secure connection for the UE accessing the 5G CN over non-3GPP access networks. Such integration enables a UE to establish multiple
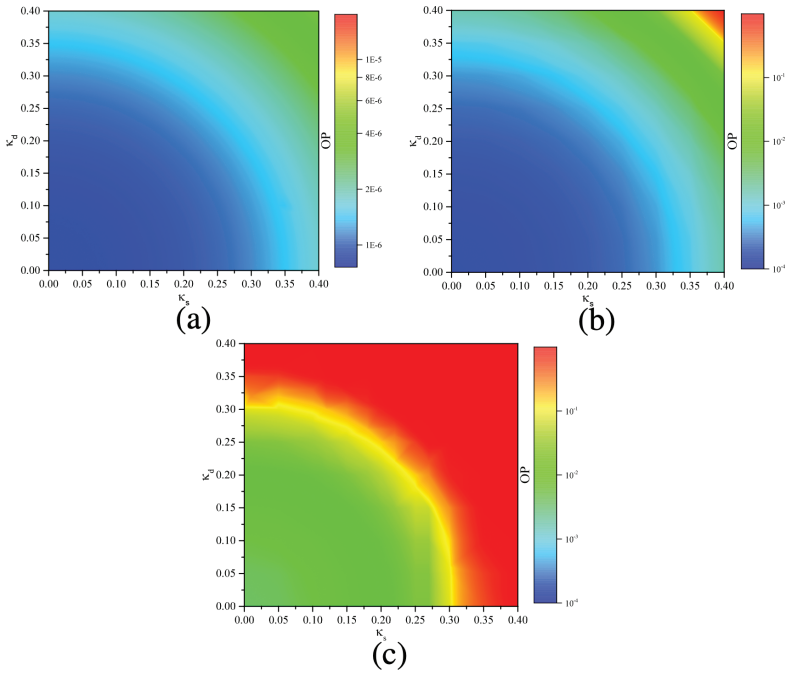
**Figure 3.25.** RIS-assisted UAV link outage probability against imperfect link parameters.

sessions to receive and send data traffic over 3GPP access and/or non-3GPP access. In other words, the UE can have simultaneous connectivity over 3GPP and non-3GPP access networks.

The procedure that steers a UE traffic onto available access networks and enables multi-access connectivity for a single UE is termed as Access Traffic Steering, Switching and Splitting (ATSSS) aka AT3S, by 3GPP [77]. AT3S framework promises to have a great importance for beyond 5G and 6G networks to enable ubiquitous connectivity and service continuity aspects by utilizing all available networks including 3GPP and non-3GPP access networks.

Current framework defined for AT3S by 3GPP has several steering modes, such as active-standby, smallest delay, throughput aggregation, load balancing, redundant, and priority based, and can use multi-path transmission control protocol (MPTCP) to send data through different available network interfaces. The widespread availability of HetNets and the enablement of multi-access connectivity from a single UE to the available access networks increase the complexity level of access traffic routing and resource management for beyond 5G and 6G networks.

An improved version of the 3GPP AT3S framework, called enhanced AT3S (eAT3S), is proposed that uses multi-wireless access technology (multi-WAT) telemetry in an xApp within the ORAN RIC framework to add near-RT RAN control of the AT3S policies behind the user plane function (UPF) [78].

Accordingly, the telemetry and performance measurement values may be used to configure how the traffic splitting shall be performed, e.g., to estimate the proportion of traffic that shall be sent over the different access traffic flows to achieve a particular objective. The implementations presented in this section are done on a 5G Wi-Fi 6 constellation; however, the learnings are relevant for beyond 5G evolution towards 6G.

The 3GPP system architecture specifies that the AT3S can support two steering functions: a high-layer steering function, based on the MPTCP protocol, and a low-layer (LL) steering function based on the AT3S function (AT3S-LL). Each steering functionality in the UE enables traffic steering, switching, and splitting across 3GPP access and non-3GPP access, in accordance with the AT3S rules provided by the network. Regarding MPTCP, the UPF may support MPTCP Proxy functionality on the network side. This capability interacts with the MPTCP functionality in the UE by using the MPTCP protocol.

Standard TCP connections are identified by a four tuple, which includes the source and destination IP addresses and source and destination ports, whose packets are sent through a single link. In the case of MPTCP, defined in RFC 6824 [79], several paths (called sub-flows) are aggregated to create one connection. The protocol includes operations to handle when and how to add or remove paths, to be compatible with legacy TCP hardware (e.g., firewalls may reject TCP connections, if the sequence numbers are not successive) and to define a fair congestion control strategy between the different links and the different hosts.

MPTCP allows the use of different packet scheduling schemes, which select the sub-flow that will forward the next packet. These schedulers may have been designed for different purposes, such as throughput aggregation, reliability improvement, and latency reduction. In particular, the MPTCP implementation available at [80] includes three main packet schedulers:

- A default MPTCP scheduler, which selects the sub-flow with the shortest smooth round-trip time (SRTT) estimated delay, which may be useful for time critical services.
- A redundant scheduler, which forwards each packet through all the sub-flows to achieve redundancy and, thus, increase reliability. Since the packet received first is processed, it also achieves low latency.
- A round-robin scheduler, which transmits packets over the different sub-flows according to a fixed cyclic schedule. Since the different sub-flows forward different data, it achieves throughput aggregation.

These concepts are implemented and can be tested in a virtualized testbed [81, 82]. A framework built on virtual machines (VMs) that incorporate numerous network interfaces as though they were the various radio access technologies has
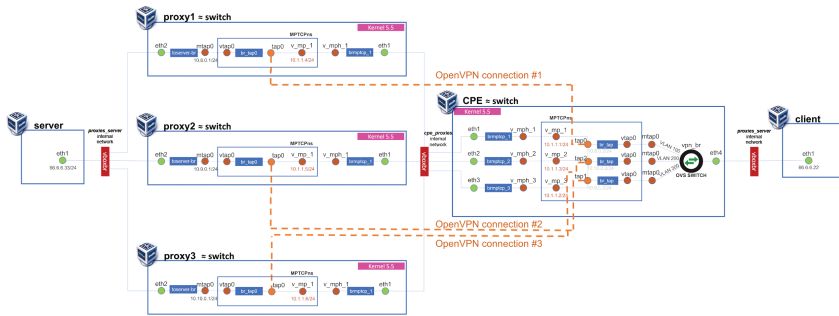
**Figure 3.26.** Virtualized multi-connectivity testbed [82].

been developed to evaluate these goals. This virtual testbed is publicly available at [82] and is depicted in Figure 3.26. In this framework, one of the VMs acts as customer premises equipment (CPE) that connects to another VM that acts as a proxy, which is located behind the 5G CN. Both the CPE and the proxy use an MPTCP-enabled kernel [80]. One (or more) client VMs attached to the CPE connect to a server behind the proxy, and a tunnel is used between the CPE and the proxy so that all traffic is transmitted using MPTCP. The main results of the evaluation of the three different schedulers are included in [78, 81]. The repository also includes a version of the testbed which connects non-3GPP networks (such as Wi-Fi 6 and LiFi) through the non-3GPP interworking function (N3IWF) using the *free5gc* platform.

Between the CPE and the MPTCP proxy, an L2 tunnel is created that enables transparent L2 traffic transit over various WATs. Since the traffic from this L2 tunnel is sent via MPTCP, all traffic passing through the CPE (such as that sent by clients connected to the CPE) is also done and is therefore scheduled using the MPTCP scheduler that has been chosen. Additionally, by establishing many L2 tunnels to various proxies, the CPE can change the packet scheduler in real time without disrupting the TCP (MPTCP) connection.

Based on this testbed, several demonstrators have been implemented [81]. Figure 3.27 shows one of these multi-connectivity testbeds that include 5G NR and Wi-Fi 6 as wireless access technologies. The adoption of MPTCP for the multi-access framework allows the performance of seamless vertical handover, since the schedulers will consider only the links that are available. The tests have shown that the data transportation also is resumed when an interface is recovered. Using these demonstrators, three aggregation modes have been tested, i.e., capacity aggregation mode, delay sensitive aggregation mode, and reliability aggregation mode [81]. These modes are described in the following subsections.

**Capacity aggregation mode:** For use cases requiring the highest user data rate, like enhanced mobile broadband (eMBB), the capacity aggregation mode might be
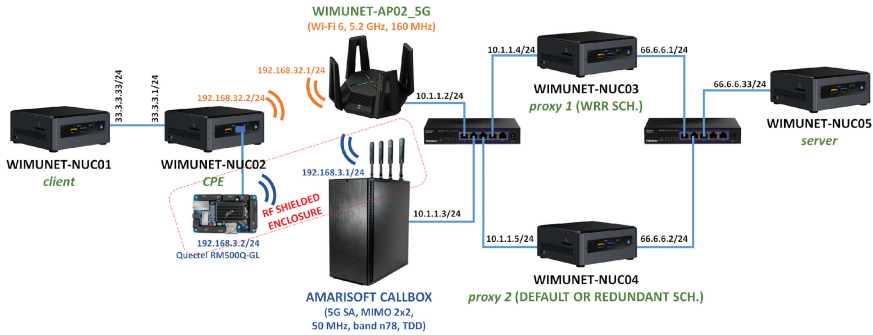
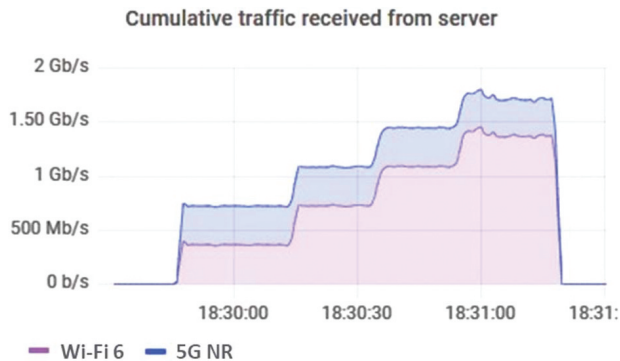**Figure 3.27.** 5G NR, Wi-Fi multi-connectivity testbed.



**Figure 3.28.** Sample results from the scheduler implemented in [81] using 5G NR and Wi-Fi 6.

beneficial. To achieve the maximum aggregate throughput, a variant of the weighted round-robin (WRR) scheduler (available at [79]) is implemented [81]. This WRR implementation allows to dynamically assign the number of turns (or weights) for each available network interface within a round.

The benefit of using this approach is twofold. On the one hand, the operator is able to adjust what proportion of traffic will be served by each radio technology. The proper assignment of the weights can lead to achieve the maximum aggregate throughput. Figure 3.28 shows an example in which the testbed reached 350 Mb/s with 5G and 1.4 Gb/s with Wi-Fi 6. As it can be observed, by using the proposed scheduler, it is possible to adjust the percentage of traffic served by each technology (varying the weights between 1–1, 1–2, 1–3, and 1–4 in this example). Moreover, with this approach, the implementation reaches the maximum aggregate throughput (around 1.75 Gb/s in this case).

**Delay sensitive aggregation mode:** For ultra-reliable and low-latency communication (URLLC) service cases, the delay sensitive aggregation mode can be beneficial.
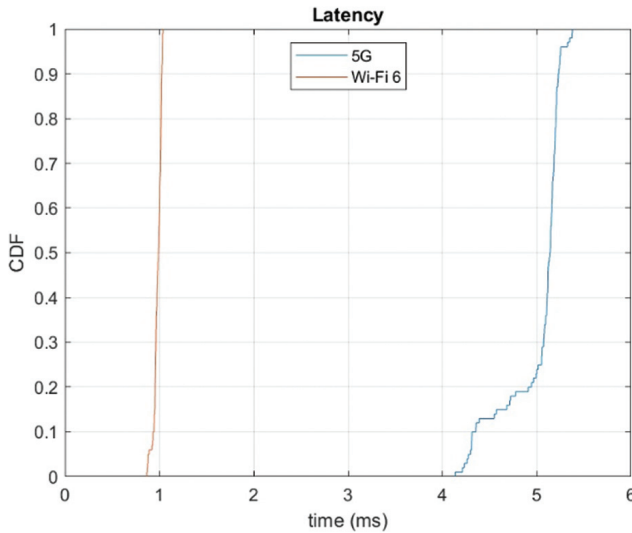
**Figure 3.29.** Latency of 5G and Wi-Fi 6 in the proposed testbed [81].

To achieve the minimum latency, the terminal shall employ an MPTCP scheduler such as the redundant (which sends the same information through all the different wireless technologies) or the default (which selects the network interface with lowest round trip time (RTT)). As shown in Figure 3.29, Wi-Fi 6 achieves a much lower latency (around 1 ms) compared to 5G (around 5 ms), mainly due to the complexity of the 5G core and the gNB (BS in 3GPP terminology), which increases the processing time.

Figure 3.30 shows the latency achieved by using multi-connectivity with both 5G and Wi-Fi 6. As expected, the latency obtained with the redundant and the default schedulers is the one achieved by the Wi-Fi 6 interface (approx. 1 ms). Even for the round-robin scheduler, the latency is reduced to approximately the average between the latency of 5G and Wi-Fi 6 (slightly higher than 2 ms). Thus, the redundant and the default MPTCP schedulers are suitable for time critical services.

These measurements have been performed using the *netperf* tool from NMAP in order to measure the latency in TCP (i.e., using MPTCP). Latency is estimated as the one-way delay, i.e., half of the RTT.

**Reliability aggregation mode:** Regarding the air interface reliability requirement, it is achieved by design because MPTCP ensures a reliable packet delivery. The trade-off introduced by MPTCP is that packets could experience too much latency, e.g., due to TCP retransmissions, thus not being useful for the application. We can, therefore, reinterpret the reliability requirement in terms of the latency requirement, which was already commented in this section.
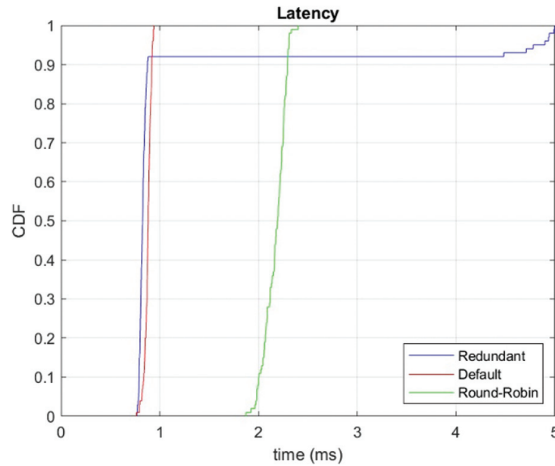
**Figure 3.30.** Latency of MPTCP using different schedulers [81].

### 3.5.1   Vertical Handover

AT3S enables simultaneous access to 3GPP and non-3GPP networks. The handover can be considered as horizontal handover or vertical handover. In horizontal handover, the UE moves within the same access network. In vertical handover, the UE moves between the 3GPP and non-3GPP access networks. When UE moves within the same access network, AT3S can be used to enable service continuity by steering the access traffic onto the other access network. Therefore, AT3S can be utilized in network performance optimizations, such as load balancing, service continuity, or seamless handover. The noted three optimization problems can also be considered as part of one single problem that focuses on the overall mobility optimization for the UE.

Within the AT3S framework, there is technically no vertical handover between access networks. In other words, there is no break-before-make approach in multi-access connectivity scenarios. Instead, when an access link is enabled, MPTCP detects the link as an available network interface and performs data traffic routing/scheduling accordingly. Therefore, the vertical handover time between available access networks can be considered as an elapsed time for an interface transition from disabled to enabled where MPTCP can start sending data through the recently enabled interface. This timer can be termed as "transmission resume delay" to mitigate any confusion on the conventionally used handover timers.

In order to measure transmission resume delay for multi-access connectivity scenarios, a controlled experiment environment is created. Accordingly, two machines are virtualized with VirtualBox version 6.1 and connected to two virtual networks via two network interfaces. Each machine runs an Ubuntu 20.4 image with the

modified kernel, which includes a specific MPTCP module designed in [81]. A test traffic is generated using the *iperf* tool. Traffic is sent from one VM (VM1) to the other VM (VM2). VM1 runs a script which detects when the network interfaces become enabled or disabled. To this end, this script uses *NETLink* sockets. This way, as soon as an interface changes, the event is captured and notified. VM1 also runs the *tcpdump* tool to capture traffic from the interface that will be enabled and disabled. Since *detect-interfaces* and *tcpdump* run in the same machine, they both use the same clock. This way, the timestamp of the event of a new available interface generated by *detect-interfaces* and the one from the *tcpdump* network capture can be used to determine the transmission resume delay. The delay is estimated by an AWK script which processes the output of *detect-interfaces* and *tcpdump*.

To simulate the outage of a link, one of the network interfaces of VM1 is disabled and enabled from the VirtualBox user interface several times during the test. A total of 100 temporal outages were simulated. As it can be seen from Figure 3.31, more than the 60% of the outages resume the transmission after 0.065 s since the interface is re-enabled. Two allocation and retention priority (ARP) resolution requests are made before sending a TCP segment, one for each destination IP address. Nevertheless, this ARP resolution lasts 0.6 ms in the worst case. The results of these experiments demonstrate that the MPTCP approach for the multi-access connectivity allows to provide continuous transmission even when a link fails. This "always-on" type of connection replaces the need for a vertical handover across access networks. Therefore, the time with no connectivity due to a handover becomes zero, and service continuity can be provided during the handover process.

Figure 3.31 shows that the measured time to resume of the transmission through a new link is nearly 60 ms. To identify the source of this delay, several experiments are conducted. To start with, Ethernet frames are forged with the source and destination MAC addresses of some of the MPTCP packets captured during the previous experiments. The payload of these frames is randomly created. A Python script that sends these frames continuously is also created. The time to resume is measured as in previous experiments. The resulting measurements show that the average delay is near to 5 ms.
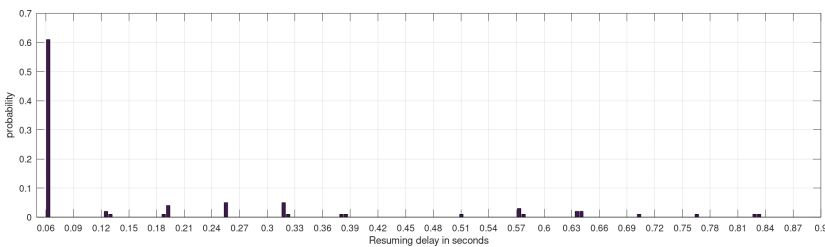


**Figure 3.31.** Time to recover from link outage [81].

## 3.6   Sub-THz for Ultra-High Data Rate

As discussed in Section 3.2.4.2, the availability of several GHz bandwidths at mmWave and Sub-THz bands makes them an attractive technology for future throughput demanding 6G use cases. Besides its potential for communication systems, the wide sub-THz spectrum is beneficial for radio sensing to increase the time, angular, and frequency resolution, in addition to providing the ability to explore the physical properties with spectroscopy [83]. However, these advantages come with challenging link budget, because of the propagation loss caused by blockage and atmospheric absorption, the decreased antenna aperture at high carrier frequencies, the increased noise associated with bandwidth, in addition to the RF hardware non-idealities and limitations. Therefore, multi-antenna implementation with high-gain beamforming is necessary to increase the link range. As a result, it is not feasible to develop one solution that fits all scenarios, but the radio design needs to be analysed for specific link requirements, considering the characteristics of the radio channel and the physical properties of the hardware components. Starting from the link requirements, in terms of data rate, range, and mobility, which are determined from the relevant use cases, the technical requirements and parameters for radio design can be analysed [84].

    This section provides an overview of the use case families subjected to the radio access in the sub-THz range (100–300 GHz), and discusses technical aspects related to link modelling, RF impairment, hardware modelling, radio, and beamforming architecture, in addition to the impact of radio channel and waveform design.

### 3.6.1   Use Cases and Technical Requirements

The relevant communication use cases require ultra-high data rate, and they can be mapped to two scenarios based on the range, which highly influence the underlying wireless technologies [5, 85]:

**Short-range wireless connectivity:** corresponds to small-cell scenario with typical cell size blow 100 m and typical peak data rate of 100 Gb/s. This scenario enables wireless access for device to infrastructure communication at a short range, for general purpose access through a hotspot deployed in indoor environment, such as home for augmented reality/virtual reality (AR/VR) applications, industrial campus for digital twin and industrial control applications, and outdoor for the applications of smart cities. This scenario can also be considered for D2D communication in the applications of digital immersion and telepresence, such as for wirelessly connecting displays and docking stations. Due to the short range of this scenario, the target mobility is low (<10 km/h).

**Long-range wireless connectivity:** can be exploited to improve coverage by providing long-range fixed wireless access to sparsely populated or hardly accessible areas, which cannot be covered by short-range infrastructure. Additionally, they are deemed a cost-effective alternative of optical fibres for the interconnection between small cells in network densification, and for backhaul. The communication range is from 200 m up to 2 km in predominantly fixed outdoor LoS environment, but it is also possible for interconnecting indoor RUs, and for providing backhaul to flying BS in non-terrestrial networks (NTN). The data rate depends on the specific scenario and can reach 1 Tb/s for infrastructure backhaul.

The basic technical requirements to achieve the peak data rate include the bandwidth, number of RF chains, modulation and coding scheme (MCS), in addition to the corresponding minimum SNR. The required SNR is derived from the link budget, which imposes transceiver design requirements including transmit power, antenna array gain, while considering the hardware impairment models, channel characteristics and mobility requirements. Moreover, the link budget is directly related to the link range [84].

### 3.6.1.1 Bandwidth and number of RF chains

The RF bandwidth $B$ (i.e., the passband bandwidth) required for supporting a peak data rate $R_p$ depends on the number of orthogonal spatial and frequency channels $M_{ch}$, the modulation order $Q_{c,m}$, and code rate $r_{c,m}$ in each chain. MCS corresponds to a number bit/symbol $L_{c,m} = r_{c,m}\log_2 Q_{c,m}$. Without considering the impact of filtering and the required guard band, and assuming the symbol rate is $1/B_m$, the relation between these parameters can be expressed as $R_p = \sum_{m=1}^{M_{ch}} L_{c,m}B_m$. Using multiple channels reduces the bandwidth requirements for the RF chain, and it is a solution when a contiguous wideband is not available. In addition, the availability of spatial channels depends on the properties of the wireless channel and the number of serving radio nodes in the case of distributed MIMO deployments. When $M_{ch}$ corresponds to the number of spatial streams, a smaller bandwidth $B = B_m$ can be used leading to an increase in the spectral efficiency, SE $= R_p/B = \sum_{m=1}^{M_{ch}} L_{c,m}$. Whereas in the case of multiple frequency channels, the aggregated bandwidth is $B_{agg} = \sum_{m=1}^{M_{ch}} B_m$, and the required spectral efficiency per channel is $SE_m = L_{c,m}$, with average SE $= \frac{1}{M_{ch}} \sum_{m=1}^{M_{ch}} SE_m$. Note that the frequency channelization does not increase the spectral efficiency in comparison to spatial multiplexing.

**LoS scenario:** assuming $B_m = B$, $r_{c,m} = r_c$, and $Q_{c,m} = Q_c$, then, $B_{agg} = M_{ch}B$, such that SE $= L_c = r_c\log_2(Q_c)$, and therefore,

$$R_p = M_{ch}B \, r_C\log_2 Q_c \Rightarrow B = \frac{R_p}{M_{ch}r_C\log_2 Q_c}. \qquad (3.1)$$

As can be seen from (3.1), there are different possibilities to choose the bandwidth and MCS parameters. For instance, when employing 16-QAM, code rate 5/6, and a single channel, the required RF bandwidth for achieving 100 Gb/s is 30 GHz, and the baseband bandwidth (i.e., the cut-off frequency of low-pass filter) is 15 GHz. This is less or equal to 10% of the carrier frequencies in the range 150–300 GHz, which is seen as the target for RF circuity design. Each MCS requires sufficient SNR at the receiver to fulfil predefined performance measure, such as bit error rate (BER) and packet error rate (PER) [LJT+19] or based on information theoretical analysis with additional margin for the implementation algorithms, as discussed in the next subsection.

### 3.6.1.2   SNR requirements

Using the additive noise model after equalization, $y = d + v$, the SNR of the symbol is defined by SNR $= \frac{E[|d|^2]}{E[|v|^2]}$. Shannon formula assumes additive white Gaussian noise (AWGN) channel and Gaussian symbols, and thus, the minimum SNR can be derived from $\log_2(1 + \text{SNR}) \geq L_c$. However, considering discrete constellation $\mathcal{M}_c = d_0, \ldots, d_{Q_c-1}$ and uniform distribution of the symbols, the mutual information for a noise variance $\sigma^2 = E[|v|^2]$ can be computed from the formula

$$I(\mathcal{M}_c, \sigma^2) = \log_2 Q_c$$

$$- \frac{1}{Q_c \pi \sigma^2} \int \sum_0^{Q_c-1} e^{-\frac{|y - d_m|^2}{\sigma^2}} \log \left( \sum_0^{Q_c-1} e^{-\frac{|y-d_m|^2}{\sigma^2}} \right) dy. \quad (3.2)$$

Figure 3.32 shows the theoretical achievable mutual information for different QAM order at different SNRs. As can be seen from the figure, for uncorded modulation, the minimum SNR for uncorded 16-QAM is about 17.5 dB so that $L_c \approx 4$ bits/symbol. By giving sufficient margin for processing, a reliable transmission can be achieved. Moreover, by employing channel coding with code rate $r_c = 1/2$, $L_c = 2$ bits/symbol. In theory, the minimum SNR for the latter case is 5 dB. However, by allowing a margin for decoding algorithms, the required SNR is higher. Practically, the SNR requirements in the coded case compared to the uncoded one can be reduced by $10 \log_{10}(r_c)$ dB. Thus, in the case of 16-QAM and $r_c = 1/2$, a minimum SNR of 14.5 is sufficient.

The SNR is calculated from the link model, and it is proportional to the ratio between the transmitted power and bandwidth, $\frac{P_{TX}}{B}$, per chain. Based on that, the required SNR in case of employing multiple frequency channel for the same MCS is identical. Since the bandwidth $B = \frac{B_{\text{agg}}}{M_{ch}}$, and the total power $P_{TX,\text{tot}} = M_{ch}P_{TX}$, the required SNR does not change with frequency channelization. Nevertheless,
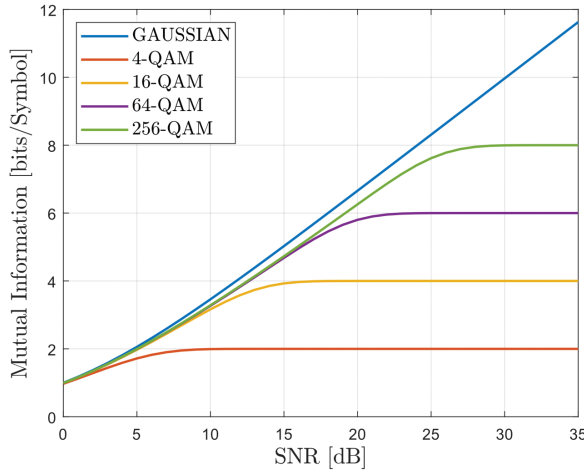
**Figure 3.32.** Mutual information for QAM.

the realization of the hardware architecture to achieve such SNR can be relaxed for smaller bandwidth.

### 3.6.1.3   Link model

The abstracted link model of the digital baseband signal relates the transmitted signal $x[n]$ to the received signal $y[n]$. In LoS, the link model can be expressed in the form

$$y[n] = hx[n] + z[n] + v[n], \tag{3.3}$$

where $h$ is the channel gain, $v[n]$ the additive noise that is independent of the signal, and the additional term $z[n]$ denotes the self-interference term because of the hardware non-idealities. The channel gain is related to the link budget, the noise power is determined from the thermal noise, the noise figure, and the bandwidth, and the interference power is computed from the hardware models. In the low SNR region, the receiver performance is inevitably noise limited, while in the higher SNRs the performance is limited by the non-linearity. Note that $x[n]$ can be the modulated signal, and thus, after demodulation, additional processing gain can be achieved. In fact, (3.3) considers the linearization of the hardware response, where the RF non-idealities are treaded as additive noise. In general, a non-linear model can be used such that

$$y[n] = f(h, x[n]) + v[n], \tag{3.4}$$

where $f(h, x[n])$ is a non-linear function that depends on the hardware models.

As illustrated in Figure 3.33, which focuses on one signal chain, the link budget, is computed from the relation between the received power $P_{RX}$ and the transmitted
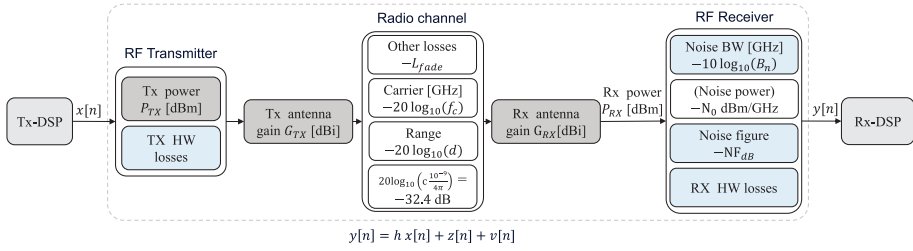
**Figure 3.33.** Link budget model.

power $P_{TX}$, given the total antenna gains $G_{a,TX}$ and $G_{a,RX}$ at the transmitter and receiver, respectively, in addition to the propagation losses, such that [85, 86]:

$$P_{RX}[dBm] = P_{TX}[dBm] + G_{a,TX}[dBi] + G_{a,RX}[dBi]$$
$$- 20\log_{10}(f_c[GHz]) - 32.4 - 20\log_{10}(d) - L_{fade}. \qquad (3.5)$$

Here, $f_c$ is the carrier frequency, $d$ is the range, and $L_{fade}$ denotes other path losses, such absorption or fading in the case of non-LoS. The link budget decreases by 6 dB by doubling the range or doubling the carrier frequency. This can be compensated, for instance, by doubling the antenna gains at the transmitter and receiver for the same transmit power.

Considering the hardware imperfections, the SNR can be expressed in the form

$$SNR[dB] = P_{RX}[dBm] - N_0[dBm/Hz]$$
$$- 10\log_{10}(B_n[Hz]) - NF[dB] - L_{HW}. \qquad (3.6)$$

The term $N_0 = 10\log_{10}(1000\,kT)$ is the noise spectral density, where $k$ is Boltzmann's constant, $T$ temperature in Kelvin (for typical temperature $T = 300$ K, $N_0 = -174$ dBm/Hz). Moreover, $B_n$ is the noise bandwidth, which can be higher than the signal bandwidth B. Furthermore, the term $L_{HW}$ represents the overall losses in the hardware. The SNR can be increased by increasing the link budget, which depends on the RF transceiver architecture. In particular, the antenna gains represent the overall antenna element gains and array gain, and $P_{TX}$ corresponds to the total power generated by the PAs. Thus, the EIRP ($P_{TX}[dBm] + G_{TX}[dBi]$) can result from the power radiated by different number of antenna elements and PAs, and the behaviour of highly directive antennas such as horn, lenses, and reflectors. The sensitivity of the receiver $P_{sens}$ is computed according to the minimum $SNR_{min}$ requirements for the targeted MCS,

$$P_{sens}[dBm] = SNR_{min}[dB] + N_0[dBm/Hz]$$
$$+ 10\log_{10}(B_n[Hz]) + NF[dB] + L_{HW}. \qquad (3.7)$$

From system level perspective, the link budget parameters highly depend on the technology, bandwidth, carrier frequency, and waveforms. For instance, the transmitted power results form

$$P_{TX} = P_{sat} - P_{BO} - L_{ant-PA}, \tag{3.8}$$

where $P_{sat}$ is the saturated (maximum) output power, $P_{BO}$ is the back off needed, and $L_{ant-PA}$ is the loss between PA and antenna. Moreover, the hardware imperfections, which contribute to $L_{HW}$, depend on parameters related to RF, and analogue and digital constraints that tend to worsen by increasing the frequency. These parameters include

- Transition frequency of the transistor $f_T$ of low-noise amplifier (LNA), which impacts the receiver noise.
- Frequency of unity unilateral gain $f_{max}$ of PA, which impacts the transmitted power.
- Parameters impact the detection including the phase noise $N_{ph}$, digital gate delay ($t_{Gate}$), clock frequency ($f_{clk}$), clock jitter $\Delta t_{jitter}$, quantization noise $N_q$, and maximum signal level $V_{max}$.
- Other signal-dependent non-idealities or distortion components $D_{other}$, and cross-coupling between channels $I_{mutual}$.

## 3.6.2 Radio Design Consideration

A functional block diagram of potential single phased array transmitter and receiver models is shown in Figure 3.34, whereas the overall system may consist of several chains. At the transmitter, the complex-valued discrete complex signal $x[n]$ is converted to baseband signals, $x_I(t)$ and $x_Q(t)$, using DAC and LP. The baseband is upconverted to RF signal $s(t)$ using mixers assuming direct conversion. The signal is spitted to parallel signals at the beamforming module that pass through the gain and phase control blocks, then to the PA, and finally to the antenna elements. The overall transmitter signal model considers the impacts of DAC quantization, I/Q imbalance, phase noise, analogue beamforming, PA non-linearity, and the antennas. The received signal at each antenna is amplified by an LNA connected to gain and phase control at the receiver beamforming block, then the output signals are compinged in the received signal $r(t)$. The baseband signals $y_I(t)$ and $y_Q(t)$ are generated by the IQ mixer and LP filter, which are converted by the ADCs to obtain the discrete complex signal $y[n]$. Similarly, the transmitter signal model contains the impacts of ADC quantization, I/Q imbalance, phase noise, and analogue beamforming. The LNA can be non-linear as well, which impacts the analogue gain control (AGC) functionality.
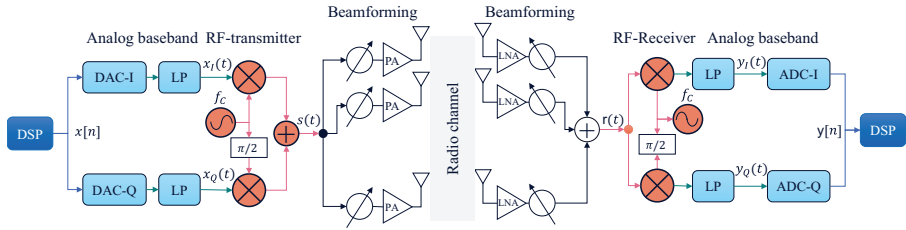
**Figure 3.34.** Functional RF blocks of a single phased-array transmitter and receiver.

Although the transceiver architecture looks similar to the one used in 5G mmWave, the key difference is in the non-ideal behaviour of the RF components at the frequency range above 100 GHz and ultra-wideband, as well as the increased technical challenges in the hardware implementation. Moreover, to achieve data rate larger than 100 Gb/s at a specified range with affordable complexity, form factor, and energy consumption, implementation with multiple RF chains is required either in the form of multiple aggregated frequency channels, in addition to the consideration of hybrid beamforming architecture depending on the availability of spatial beams. Therefore, several crucial aspects need to be considered in the radio design:

**Duplex mode:** TDD operation is envisioned because of the highly integrated modules of the antennas and the related LNAs and PAs as well as a large number of antennas. TDD allows also exploiting radio channel reciprocity as the same band is used for transmission and reception. However, to exploit this feature, the transmitter and receiver need to be calibrated to account for different frontend hardware responses. In addition, the antenna can be used for transition and reception, which requires a T/R switch at the cost of additional losses between PA and the antennas [87]. To avoid such losses, another option is to use separate antenna arrays for the transmitter and receiver. In this case, the channel reciprocity is not fulfilled, but still the TDD operation is favourable to avoid costly diplexers [88].

**Transmit power:** Generating sufficient power is one of the most important and well-known challenges in wireless communications, which become more challenging with the increase of frequency. In RF performance, the power is limited by the capabilities of transistors implemented using different semiconductor technologies, and as a general trend the output saturated power decades severely as a function of frequency, especially, above 100 GHz. As this power corresponds to non-linear operation, sufficient back-off needs to be considered, which depends on the waveform and increases consumption [89].

**Receiver noise:** Like any other electronic devices, both active and passive receiver components contribute to the receiver noise, which consequently degrades the signal quality in terms of, e.g., SNR. These include insertion losses (ILs) in connected ports of different components, losses caused by interconnection and routing, in addition to the noise factor of LNA and other blocks. ILs increase at higher frequencies, which require high level of integration to mitigate. Also, the noise figure of LNA, which is the dominant active component, generally increases with the frequencies above 100 GHz, but can be less severe in some technologies [90]. In addition, as the noise is proportional to the bandwidth, doubling the bandwidth results in 3 dB losses of SNR.

**Phase noise:** It causes significant degradation in the performance of high data rate by causing random rotation of the received signal constellation and causes interference to adjacent channel in case of frequency division systems. This effect highly depends on the waveform, bandwidth, and oscillator source selection. The larger is the bandwidth, the higher is the amount of accumulated phase error, and such effect cannot simply be compensated by increasing SNR, as the noise floor becomes dominant. The phase noise depends on the technology of local oscillators (LOs). For instance, one common technique generates high frequency by multiplication of low-frequency oscillator. Each doubling of the frequency results in 6 dB higher phase noise level. Thus, it is critical to afford LO with low phase noise in order to allow operation at ultra-wideband [91].

**DAC/ADC:** The speed of converters is proportional to the bandwidth assuming Nyquist sampling. The essential challenge for high-speed converters is the increase of energy consumption with the sampling frequency and resolution. Nevertheless, the implementation of high speed DAC is relatively easier in comparison to ADC, and its power consumption at the transmitter is not significant when compared to the PA. In relation to ADC, the sampling frequency $f_s$ needs to be at least twice the baseband bandwidth to avoid aliasing, and up-sampling might be required for digital synchronization. The quantization noise should not have a big impact on the SNR if the used amount of bits is sufficient. In particular, after quantization $\text{SNR}_d = \frac{P_S}{\sigma^2 + \sigma_q^2} = \frac{P_S}{\sigma^2} \frac{1}{1 + \sigma_q^2/\sigma^2}$, where $\sigma_q^2$ is the quantization noise power, if $\sigma_q^2/\sigma^2 < 0.1$, the reduction of the SNR $\frac{P_S}{\sigma^2}$ is less than 0.4 dB. Therefor, the ADC dynamic range $\text{SNR}_{\text{ADC}}$ should be considered according to the minimum SNR, such that, $\text{SNR}_{\text{ADC}}[\text{dB}] > \text{SNR}_{\text{min}}[\text{dB}] + 10$. In uniform quantization, $\text{SNR}_{\text{ADC}}[dB] = 6.02 Q_d$, where $Q_d$ is the resolution in number of bits. Moreover, to avoid clipping, compensate for AGC, and tackle the near-far problem at the BS, additional margins are required. All that increases the required resolution, which increases the ADC power consumption $P_{\text{ADC}}$ according to Walden's figure-of-merit

(FoM), which is used to evaluate ADC performances and given by [92].

$$\text{FOM}_w = \frac{P_{\text{ADC}}}{2^{Q_d} \cdot f_s}. \tag{3.9}$$

This parameter depends on the technology, and it is shown to be constant up to 100 MHz, where doubling the sampling rate or increasing the resolution by 1 bit double the power consumption. In contrast, for sampling rate above 1 GHz, $\text{FOM}_w$ significantly increases as a function of the sampling rate, and therefore the power consumption grows more than double by doubling the sampling rate.

**DSP:** The implementation of real-time signal processing algorithms to achieve reliable throughput larger than 100 Gb/s within low-latency constrained is challenging and requires processing papalism, which leads to high energy consumption as well. Therefore, the DSP architecture needs to be optimized, such as considering the design of low-complexity algorithms without significant performance degradation and implementation with low-resolution operations. This might require adding an additional SNR margin, which impacts the RF design, and thus, a joint optimization of RF hardware and DSP is important [93].

**Antenna:** As the wavelength decreases, the antenna sizes also decrease, which create challenges for the transceiver radio-frequency integrated circuit (RFIC) design to match the form factor of the electronics and the antenna matrix [94]. For instance, to increase the transmit power, the size of PA consumes larger area than the antenna, which limits the practical size of the antenna array. Nevertheless, an array of large number of elements is required to increase the aperture and compensate for the reduced physical size of a single antenna. The array gain, assuming $\lambda/2$ element spacing, is proportional to the number of elements $N$, $G_{a,\,x} = NG_x$, where $G_x$ is the antenna element gain. Following the link budget model (3.5), doubling the frequency or doubling the range requires doubling the number of elements at the transmitter and receiver sides, or to increase by factor 4 at one end of the link. Thus, employing antenna array at frequencies above 100 GHz, for sufficient range, is challenging. Besides the design RFIC design and packaging limitations, antenna arrays are impacted by non-idealities such as coupling between antenna elements. An alternative option of large antenna array is using high directivity antenna such as horn and lenses, but this is not steering-friendly choice [95].

**Beamforming:** The high-gain antenna arrays required to improve the link budget produce narrow beams, which requires strict alignment between the transmitter and receiver. Moreover, upon mobility, switching the beam needs to be fast enough when analogue beamforming is employed. However, because of the limited angular

resolution because of the limitation of the gain and phase control, a successful beam alignment might not be guaranteed in all positions. Although digital beamforming is fully flexible, it is not practically feasible, and thus, hybrid beamforming arises as a reasonable approach. In hybrid beamforming, the antenna array is partitioned into subarrays driven by multiple RF chains. This allows also serving multiple users or exploiting MIMO spatial multiplexing. The size and number of subarrays are design parameters to be considered in the radio design based on the scenario and channel characteristics. Other challenges for the beamforming design are inherited from the non-idealities of other RF components, the constrained on the transmitted power limits the exploitation of gain control, the violation of the narrow band approximation, violation of the far-field assumption with large size of the array relative to the wavelength, and the antenna coupling.

### 3.6.3   RF Hardware Modelling

The modelling of individual block is essential for the computation of the link budget and the design of the waveform to mitigate the non-linearity in the system, in addition to be able to estimate the performance of realistic system. Three modelling approaches are required based on the studied problem, namely

- **Second order statistics:** this model considers the second-order statistics only, to compute the SNR ($P_{RX} = P_{TX} + P_N$) for link-level range and coverage analysis. This modelling can be carried out to determine the contribution of each block and identify the dominant source of non-ideality.
- **Additive non-linearity model:** in this approach, the RF impairment is considered as an additive term after linearization according the signal model $y[n] = hx[n] + z[n] + v[n]$, and $h$ is the linearization gain, and $z[n]$ is signal-dependent term corresponds to the overall all RF block non-idealities such as quantization noise, phase noise, non-linear distortion of PA, IQ imbalance, etc. Thus, $z[n]$ can be approximately modelled by simplified Gaussian noise considering the sum of different random non-idealities. This model is useful to perform simulation in affordable time to evaluate the link performance in terms of BER, FER, etc., but it is not necessarily realistic due to the non-Gaussian nature of the signal-dependent RF non-idealities.
- **Response function:** where the goal is to obtain a function representing the relation between the input and output of blocks such as PA and LO to study compensation techniques. These models are mostly memory-dependent, such that the current response depends on the previous state. The simulation using this model is time consuming, as it emulates the actual hardware. The overall response of the system can be expressed as $y[n] = f(x[n], x[n-1], \ldots) + v[n]$.

The following subsections provide an overview of modelling examples for different components.

### 3.6.3.1  ADC/DAC models

DAC and ADC add impairments to the signal by non-linear clipping and quantization noise, in addition to the effect of anti-aliasing and reconstruction filtering, as illustrated in Figure 3.35.

The filtering effects can be neglected by designing the signal to occupy a spectrum in the flat region of these filters [96]. Thus, it is required to model the impact of the quantization, which involves proper scaling of the signal at the ADC input, and proper mapping of floating point or fixed-point representation at the DAC input. The scaling of the input signal is achieved by AGC at the receiver, whereas the scaling of discrete signal is achieved digitally. Two models are shown in Figure 3.36, the first corresponds to the response function, where the signal is scaled and quantized, whereas the other model performs scaling and adds quantization noise, which can be modelled as uniformly or Gaussian-distributed random signal.

### 3.6.3.2  Power amplifier

In addition to the maximum saturated power of PA, modelling the non-linearities and memory effects are important. Measured data are used to model the PA response in the discrete-time domain. Following memory polynomial model, which
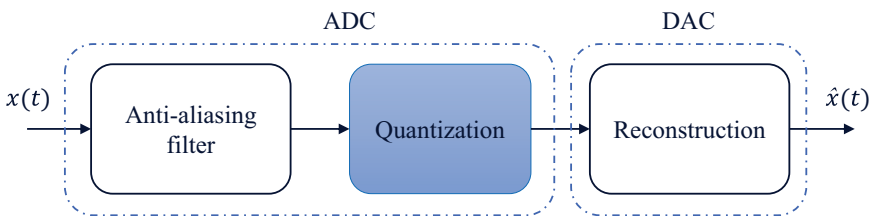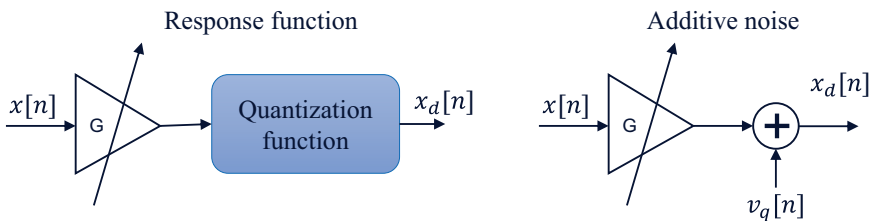


**Figure 3.35.** ADC/DAC functional model.



**Figure 3.36.** ADC/DAC signal models.

can be expressed as [97]:

$$V\text{out}(s) = \sum_{q=1}^{Q} \int_{k=1}^{K} \tilde{a}_{kq} V\text{in}(s-q)|V\text{in}(s-q)|^{2(k-1)}, \quad (3.10)$$

where $s$ is the number of samples, $K$ polynomial order, $Q$ memory length, and $\tilde{a}_{kq}$ are the coefficient to be estimated by fitting. The signals $V_{in}$ (s) and $V_{out}$ (s) are the measured discrete input and output complex envelope signals of the $s$-th sample, respectively. Note that the selection of the proper PA model depends highly on the scenario that is investigated. For example, for linearization studies, very accurate models are required, while for generic link-budget investigations, rather simplified behavioural approaches may be sufficient to achieve the correct order of magnitude for the modelling.

### 3.6.3.3   LNA noise figure

The losses and noise contribution from components after the LNA are mitigated by the power gain of the LNA, and the major noise impact is introduced by LNA. The noise-figure of LNA depends on the frequency of operation, which can be estimated based on measured data of certain semiconductor technology or based on circuit simulation. Based on empirical results and curve fitting in the mmWave range and sub-THz range (30–300 GHz), a common noticeable trend among the investigated technologies is the exponential increase of the noise figure with frequency [98], which is defined by exponential function parametrized by two parameters $\alpha$ and $\beta$, which depends on the technology, such that the minimum noise figure, $NF$, is given by

$$NF_{\min} = \alpha \, exp\left(\frac{f_c}{\beta}\right) \quad (3.11)$$

where exemplary values are provided in Table 3.4.

**Table 3.4.** LNA noise minimum noise figure parameters of (3.11).

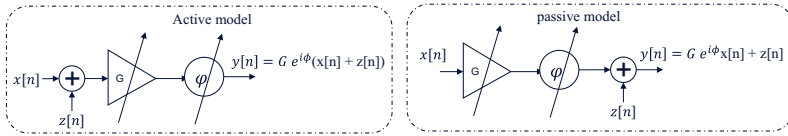|          | CMOS  | SiGe  | GaAS  | InP   |
|----------|-------|-------|-------|-------|
| $\alpha$ | 1.50  | 1.75  | 0.70  | 1.50  |
| $\beta$  | 112.4 | 130.7 | 129.9 | 188.7 |

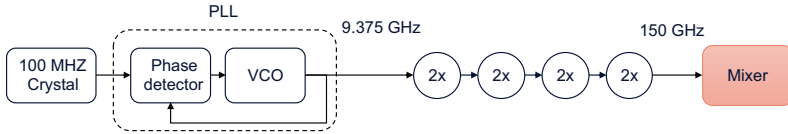**Figure 3.37.** Analogue phase shifter models.



**Figure 3.38.** Example of LO generation architecture to support 150 GHz frequency.

### 3.6.3.4 Analogue beamforming non-linearities

The gain and phase control block, which can be implemented by different passive and active technologies, is inserted before the PA at the transmitter and after the LNA at the receiver. The non-idealities of this RF block result from the quantized beamforming weights and potential noise that has impact, especially if the amplitude is controlled. In most cases, this impact is neglected at the receiver compared to the impact of LNA noise, as sufficient gain is assumed from the amplifier. However, when the amplitude is controlled for beamforming purpose, modelling the noise of the VGA/phase shifter should be considered at least to analyse a realistic amplitude control dynamic range. A simplified modelling considers additive noise, as shown in Figure 3.37, which can be added at the input in the case of active component, or after in the case of passive components. At the transmitter side, the noise can be neglected in the overall link model. The transfer function model can be derived from the employed components for the evaluation of the beamforming.

### 3.6.3.5 Phase noise

The total phase noise highly depends on the frequency synthesis approach of the transceiver. By utilizing multiple multiplications by two or three of the voltage-controlled oscillator (VCO) signal in series, as shown in Figure 3.38, where the operation frequency of the VCO can be lowered for practical implementation [99]. However, doubling the LO frequency effectively raises the respective phase noise by 6 dB since the same signal is in intermediate frequency (IF) and LO port of the mixer and correlated noises are multiplied into the RF port of the mixer.

The single-side phase noise power spectrum can be defined by [100], Section 4.2.3.1.

$$S(f_o) = \text{PSD0} \frac{\prod_{n=1}^{N} \left(1 + \left(\frac{f_o}{f_{z,n}}\right)^{\alpha_{z,n}}\right)}{\prod_{m=1}^{M} \left(1 + \left(\frac{f_o}{f_{p,m}}\right)^{\alpha_{p,m}}\right)}, \tag{3.12}$$

where $f_o$ is the offset frequency, $\{f_{z,n}\}$ zeroes of orders $\{\alpha_{z,n}\}$, $\{f_{p,m}\}$ poles of orders $\{\alpha_{p,n}\}$. The average root mean square phase jitter, which can be further converted to an equivalent SNR or error vector magnitude (EVM), can be used in the simulation, and it can be defined as

$$\sigma_{\text{rms}} = \sqrt{2 \int_{f_1}^{f_2} S(f) df}. \tag{3.13}$$

The integration goes from minimum to maximum frequencies that matters, such as the frequency range of the signal over which the phase compensation is performed. Accordingly, a signal model using random phase $\varphi_{pn}(n) \sim N(0, \sigma_{\text{rms}}^2)$ can be used, such that

$$y(n) = x(n) \exp(j\varphi_{pn}(n)). \tag{3.14}$$

This is suitable for short symbols, but to elaborate the impact of longer signal, a better model can be obtained by filtering Gaussian noise whose frequency response is $\varphi_w(f)$ using the filter $H(f) = \frac{1}{\sqrt{2}}\sqrt{S(f)}$, such that the generated phase noise is given by

$$\varphi_{LO}(t) = \int_{f_1}^{f_2} H(f)\varphi_w(f)\exp(2\pi ft) df. \tag{3.15}$$

### 3.6.4   Radio Architecture

A generic block diagram of the radio architecture that employs hybrid beamforming is illustrated in Figure 3.39. The transmitter is equipped with $M$ RF chains and a phased antenna array with $P > M$ antenna elements, which are partitioned into $M$ subarrays. A total number of antenna elements is $Q > N$ split into $N$ subarrays connected to $N$ receive RF chains at the receiver. A single RF chain follows the diagram presented in Figure 3.34. This generic architecture shows various degrees of
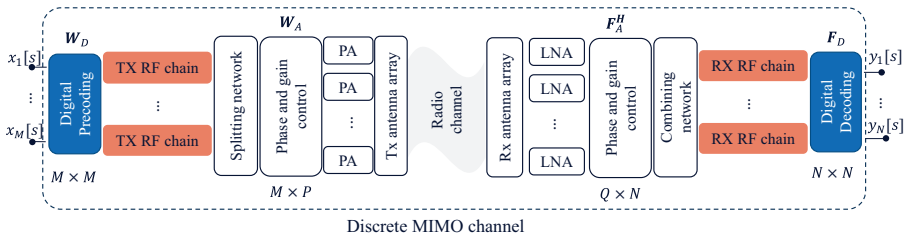


Figure 3.39. Generic hybrid beamforming radio architecture.

freedom that need to be determined at the design stage, and also at the operation, depending on the scenario and requirements. The defined configuration should provide the suitable beamforming architecture, based on the radio channel characteristics, in addition to the waveform properties before and after digital precoding, if applicable, to cope with the RF hardware limitations. The following subsections introduce an overview of the relevant radio modules.

### 3.6.4.1 Beamforming

The beamforming is achieved by means of digital precoding and analogue beamforming. The beamforming at the transmitter is achieved by a digital precoding matrix $\boldsymbol{W}_D$ of size $M \times M$ and analogue beamforming matrix $\boldsymbol{W}_A$ of size $M \times P$. Similarly, the receiver employs analogue beamforming matrix $\boldsymbol{F}_A$ of size $N \times Q$ and digital decoding matrix $\boldsymbol{F}_D$ of size $N \times N$.

The effective MIMO signal model, assuming ideal hardware is given in the frequency domain by

$$\widetilde{\boldsymbol{y}}(f) = \boldsymbol{F}^H \widetilde{\boldsymbol{H}}(f) \boldsymbol{W} \widetilde{\boldsymbol{x}}(f) + \boldsymbol{F}^T \widetilde{\boldsymbol{v}}(f), \qquad (3.16)$$

where $\widetilde{\boldsymbol{H}}(f)$ is the $M \times M$ MIMO channel in the frequency domain, and $\boldsymbol{W} = \boldsymbol{W}_A \boldsymbol{W}_D$ and $\boldsymbol{F} = \boldsymbol{F}_A \boldsymbol{F}_D$. This model can be used for initial assessment of the theoretical limits w.r.t. the channel characteristics. For instance, to select the number of RF chains based on the rank of the channel, evaluate the achievable rate and compare the performance of initial beam access techniques. Moreover, from implementation perspective, the structures of $\boldsymbol{W}_A$ and $\boldsymbol{F}_A$ are constrained by the implementation of the beamforming. For instance, as in practice, each antenna element is connected to phase and gain control, $\boldsymbol{W}_A$ is sparse, where non-zero values appear only once in each column and row. Moreover, the non-zero entries are limited by the resolution of the beamforming control.

### 3.6.4.2 Waveform and precoding

The design of the waveform and precoding needs to consider mitigating the hardware limitations at frequency above 100 GHz. This includes robustness to the phase noise as a major impairment that cannot be solved without compensation even at high SNR. To reduce the back-off and, thus, allowing higher transmit power, the waveform needs to be chosen to have smaller peak-to-average power ratio (PAPR). Moreover, to reduce the required ADC power, it is important to consider waveforms that work with low-resolution quantization. Furthermore, the processing complexity, such as equalization and detection, needs to be minimized to reduce the power consumption by DSP, especially when targeting data rate above 100 Gb/s. In addition, considering the channel characteristics, and the impact of beamforming architecture, it is foreseen that 6G waveform might need to go beyond orthogonal

frequency-division multiplexing (OFDM). For example:

**Zero-Crossing Modulation (ZXM):** which is based on temporal oversampling and 1-bit quantization. The information bits are encoded in the zero crossing of the signal. Although it offers low complexity and power efficiency in AWGN, its performance under phase noise and in selective channel needs further investigation [101].

**Analogue Multicarrier:** is based on dividing the wideband to multiple narrower bands, where each band is modulated independently or considering overlap with other bands, in similar fashion to analogue OFDM. This allows using conventional wideband waveforms per sub-band at lower frequencies, then aggregating the bands at the high frequency. This approach allows using high-resolution ADC without linear scale of the ADC power consumption, but the impact of the high-frequency phase noise needs to be evaluated [102].

**DFTS-OFDM and SC-FDE:** are preceded version of OFDM using discrete Fourier transform (DFT) matrix, which is smaller than the DFT size in the first, and the same as the DFT in the later, with the goal of achieving a trade-off between reducing the PAPR and complexity. The impact of phase noise and low-resolution ADC should be considered.

Other waveforms include constant phase modulation (CPM) and its constrained envelope with a main focus on reducing the PAPR by employing precoding on OFDM [103].

## 3.6.5 Radio Channel

The radio channel characteristic affects several KPIs, and the path loss impacts the SNR and thus the data rate. Rich multi-path environment and polarization allow exploiting spatial multiplexing and thus increasing the spectral efficiency. The spatial correlation affects the area traffic capacity and connection density. The channel dynamics such as fading in addition to the path loss has a significant impact on the energy efficiency, in terms of transmitted power and signal processing for channel estimation and equalization. This dynamicity has also an impact on the latency and reliability as going in deep fade results in lost packets and a need for retransmissions. The impact of the channel needs to be studied in terms of wave-material interaction, atmospheric losses, and multi-path characteristics. In addition, to the propagation models, which are independent of the hardware, the channel model should consider the number of antennas and the RF design parameters, in order to provide a discrete channel model suitable for designing the signal processing algorithms. This methodology is conventional in all wireless communications systems. However, it becomes challenging for frequencies above 100 GHz, because

of the complexity of designing a channel sounder and obtaining measurements in different environments.

### 3.6.5.1  Path loss, angular, and delay dispersions

The path loss, delay, and angular spread models of a wireless channel are useful information to predict the characteristics of the channel. These parameters are estimated for the LoS and NLoS links in the entrance hall and three outdoor scenarios, namely suburban, residential, and city centre, described in [86].

Two path loss models, such as the close-in (CI) reference free space reference and the alpha-beta-gamma (ABG) models, are considered here. The path loss of a signal with frequency $f_c$ at distance $d$ based on the CI model is expressed as

$$PL^{CI} = \text{FSPL}(f_c, 1 \text{ m}) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma^{CI} \tag{3.17}$$

where $\text{FSPL}(f_c, 1 \text{ m})$ is the free space path loss of the signal with frequency $f_c$ at 1 m distance, $n$ is the path loss exponent, and $\chi_\sigma^{CI}$ is the CI shadow fading. Meanwhile, the ABG model is given by

$$PL^{ABG} = 10\alpha\log_{10}(d) + \beta + 10\gamma\log_{10}\left(\frac{f}{1 \text{ GHz}}\right) + \chi_\sigma^{ABG} \tag{3.18}$$

where $\alpha$ is the distance-dependent loss coefficient, $\beta$ is an offset coefficient, $\gamma$ is the frequency-dependent loss coefficient, and $\chi_\sigma^{ABG}$ is the ABG shadow fading. The omnidirectional path losses for the indoor and outdoor scenarios, together with the fitted path loss models, are plotted in Figures 3.40 and 3.41.
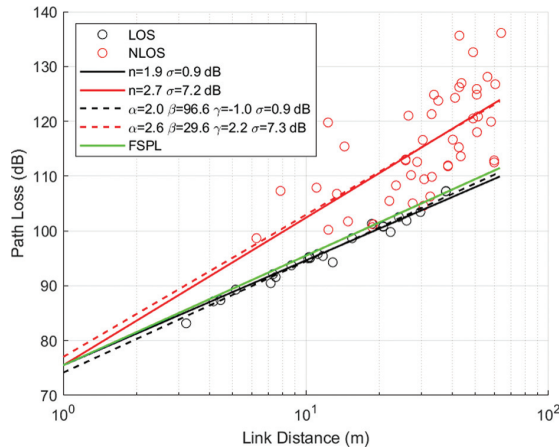


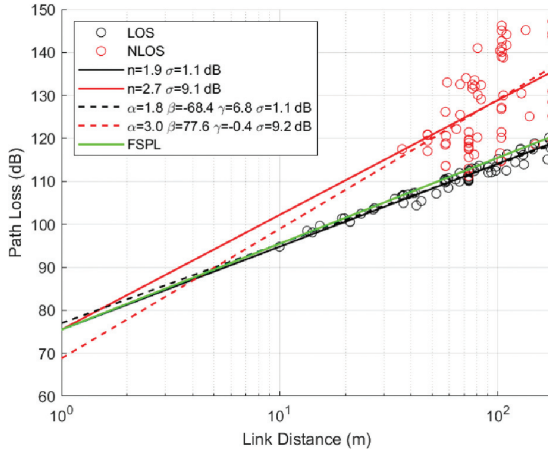**Figure 3.40.** Path loss model for indoor scenario.

**Figure 3.41.** Path loss model for outdoor scenario.

Note that the outdoor scenario is an ensemble of path loss estimates from the three mentioned outdoor scenarios, since there are only limited NLoS links that can provide sensible fitting to the model when individual scenario is considered. It can be noticed that both the CI and ABG models provide equally good fit for LoS and NLoS links in both scenarios.

The mean $\mu$ and standard deviation $\sigma$ of angular and delay spread values are listed in Table 3.5. The angular spread values are based on azimuth and zenith angle of departure (AoD/ZoD) and azimuth and zenith angle of arrival (AoA/ZoA) estimated from measurement-based ray launcher presented in [104]. Similar to the path loss modelling, the spread statistics for the outdoor case include the spread values from the three outdoor scenarios.

### 3.6.5.2   Stored channel model

The use of measured channel responses for PHY design and evaluation on a computer has been a well-recognized approach, e.g., [105], as it allows repeatable tests and comparison between different PHY schemes. The fact that measured channel responses serve as the ground-truth of any simulation-based channel modelling also justifies the use for realistic evaluation of any radio systems. There are, however, also challenges of using measured channels for PHY studies, i.e., (a) it is not straightforward to apply the measured channel responses to simulations that assume different hardware requirement than the measurement, e.g., signal dynamic range, antennas/arrays, system bandwidth, and moving speed of a mobile; and (b) a sufficient amount of measurements, e.g., for Monte-Carlo simulations and evaluation of packet error rates, may not be available due to limited capability of channel sounding. These challenges are overcome, e.g., (i) by under-sampling or interpolating the measured channels in space, bandwidth and time; and (ii) by making a general

**Table 3.5.** Angular and delay spread statistics in indoor and outdoor scenarios.

| | | AoD [deg] | | ZoD [deg] | | AoA [deg] | | ZoA [deg] | | Delay Spread [ns] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Scenario | | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ | $\mu$ | $\sigma$ |
| **Indoor** | **LoS** | 25 | 9 | 6 | 5 | 14 | 8 | 3 | 2 | 14.6 | 4.4 |
| | **NLoS** | 38 | 13 | 8 | 5 | 22 | 11 | 4 | 2 | 26.3 | 11.8 |
| **Suburban** | **LoS** | 10 | 7 | 2 | 1 | 9 | 10 | 1 | 1 | 25.7 | 24.2 |
| | **NLoS** | 9 | 11 | 3 | 3 | 5 | 4 | 1 | 2 | 15.1 | 14.8 |
| **Residential** | **LoS** | 13 | 9 | 2 | 2 | 11 | 10 | 1 | 1 | 24.9 | 19.4 |
| | **NLoS** | 20 | 18 | 3 | 3 | 6 | 6 | 1 | 2 | 26.9 | 37.9 |
| **City Centre** | **LoS** | 18 | 9 | 4 | 6 | 13 | 4 | 2 | 1 | 21.3 | 9 |
| | **NLoS** | 24 | 17 | 4 | 3 | 14 | 9 | 2 | 2 | 25.4 | 19.9 |
| **Outdoor** | **LoS** | 12 | 9 | 2 | 3 | 10 | 9 | 1 | 1 | 24.7 | 20.6 |
| | **NLoS** | 21 | 18 | 3 | 3 | 9 | 8 | 2 | 2 | 25.6 | 31.0 |
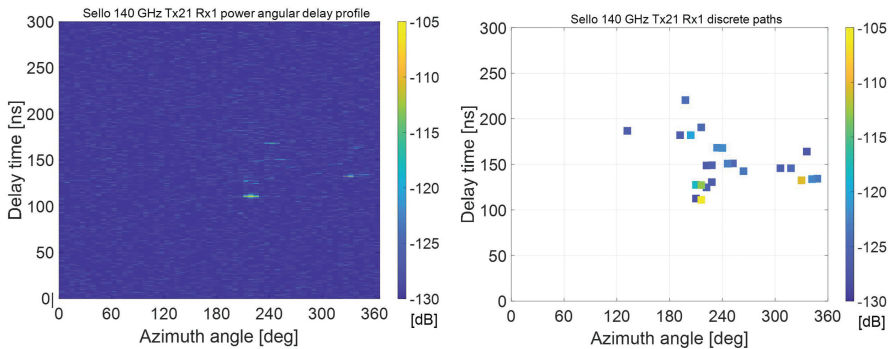


**Figure 3.42.** (a) Band- and aperture-limited channel response from a measurement and (b) its band- and aperture-unlimited model as propagation paths. Data are from a shopping mall measurement at 140 GHz [107].

mathematical description that allows us to over-sample, extrapolate, or invent the measured reality, which is called a channel model. The approach (ii) is exemplified in [105], where band-/aperture-/time-sampling-limited measurements of channels are approximated by the band-/aperture-/time-sampling unlimited form as illustrated in Figure 3.42. The former, coming from calibrated measurements, is represented by power spectrum, while the latter is a discrete model of propagation paths represented by Dirac delta functions. It is possible to synthesize infinite amount of small-scale fading realizations from the propagation paths of channels by applying

the uniform randomly distributed phases to each path before summing them up at the antenna.

The measured sub-THz multi-path channels in the discrete path format are available in [106] for an entrance hall and outdoor scenarios, such as suburban, residential, and city centre. Details of the measurement channel sounder and measurement sites are available in [86]. Even though the measurements are single-directional to cover only one link end to estimate multi-path angles, double-directional data were derived by exploiting the available detailed geometric database of the measurement environment and measured channel response using a tool called a measurement-based ray launcher [104]. These single-directional and double-directional multi-path data, in the form of discrete multi-paths, are published in [105].

Stored multi-path data cover several transmit and receive locations in different environments. Each link, i.e., transmit and receive location, is characterized by a band-/aperture-unlimited double-directional power angular delay profile (PADP)

$$P_q(\Omega^{\text{rx}}, \Omega^{\text{tx}}, \tau) = \sum_{l=1}^{L_q} P_{l,q}\delta(\Omega^{\text{rx}} - \Omega^{\text{rx}}_{l,q})\delta(\Omega^{\text{tx}} - \Omega^{\text{tx}}_{l,q})\delta(\tau - \tau_{l,q}) \quad (3.19)$$

where $q$ is the link index, $L_q$ is the number of paths, $\delta(\cdot)$ is the delta function, $P_{l,q}$, $\Omega^{\text{rx}}_{l,q}$, $\Omega^{\text{tx}}_{l,q}$, and $\tau_{l,q}$ are the power (squared magnitude of path gain), the direction of arrival, the direction of departure, and the propagation delay of the $l$th path, respectively. Arrival and departure directions contain both azimuth and elevation angles. Current data are measured with vertically polarized antennas; hence, we restrict the definitions here to a single-polarized case only, neglecting the polarization characteristics.

Receiver and transmitter antennas can be specified by complex radiation patterns $\mathbf{g}_{\text{rx}}(\Omega^{\text{rx}}) \in \mathbb{C}^{M \times 1}$ and $\mathbf{g}_{\text{tx}}(\Omega^{\text{tx}}_l) \in \mathbb{C}^{N \times 1}$, respectively, where $M$ is the number of Rx antennas and $N$ is the number of transmitter antennas. Now, the channel frequency response matrix is determined as

$$\mathbf{H}_q(f) = \sum_{l=1}^{L_q} \mathbf{g}_{\text{rx}}(\Omega^{\text{rx}}_{l,q})\sqrt{P_{l,q}}\, e^{-j2\pi f \tau_{l,q}}\mathbf{g}_{\text{tx}}(\Omega^{\text{tx}}_{l,q})^T \in \mathbb{C}^{M \times N}. \quad (3.20)$$

Random snapshots of frequency response matrices can be generated by introducing random initial phase $\theta_{l,q}$ for each multi-path, where phase terms are drawn from the uniform distribution in $[0, 2\pi]$. Moreover, the temporal dimension and time variability can be included by introducing small Doppler frequencies $\nu_{l,q}$ for each path. This models a small-scale virtual motion, where only phases of path component change over time, but other propagation parameters remain constant.

The resulting snapshot/time variant frequency response matrix is

$$\mathbf{H}_q(t,f) = \sum_{l=1}^{L_q} \mathbf{g}_{\mathrm{rx}}(\mathbf{\Omega}_{l,q}^{\mathrm{rx}}) \sqrt{P_{l,q}} \, e^{j(\theta_{l,q}+2\pi \nu_{l,q} t)} \, e^{-j2\pi f \tau_{l,q}} \, \mathbf{g}_{\mathrm{tx}}(\mathbf{\Omega}_{l,q}^{\mathrm{tx}})^T \in \mathbb{C}^{M \times N}.$$

(3.21)

### 3.6.5.3   The number of independent beams

Wireless communication over sub-THz radio frequencies demands high-gain antennas to compensate for the high propagation loss. This leads to very directive antenna patterns, which illuminate only sub-sets among all available propagation pathways. Communication systems operating at lower frequencies have extensively used spatial multiplexing and beamforming to optimally utilize all degrees of freedom provided by the propagation channel. Now, at sub-THz, partly due to the channel sparsity and mainly due to foreseen RF technology limitation, such flexible transmission schemes might not be possible. Due to aforementioned reasons, it is interesting to study how many independent beams of practical beamwidth the propagation channel support. Directional wideband propagation measurements mentioned in Section 3.6.5.2 are used for this study. One can rather easily estimate how many significant paths are present in a measurement location, but interpreting that to separable beams is not evident.

Three methods to assess the number of useful beam directions are introduced in [108]. The second method is based on measured single-directional PADPs $P_q(\mathbf{\Omega}, \tau)$ and a synthetic beam pattern $G(\mathbf{\Omega})$ defined in [109]. Only vertically polarized antennas were used in the channel measurement, and hence only single polarization is considered in the following number of beams.

Measured PADPs were evaluated using $10°$ half power bandwidth (HPBW), 10 dB dynamic range, 2 GHz BW, and 0.5 correlation threshold. An example PADP, beam power, and identified beam directions are illustrated by blue circles, a red curve, and orange squares, respectively, in Figure 3.43. Identified beam numbers in 132 measured indoor Tx and Rx locations are shown in Figure 3.44. Figure 3.45 depicts empirical CDFs of beam numbers in 132 indoor and 157 outdoor locations using both 10 and 20 dB dynamic ranges. Median values of beam numbers are two in both environments using a dynamic range of 10 dB, and using a 20 dB range, they are five and three in indoor and outdoor environments, respectively.

### 3.6.5.4   Comparison with below-100 GHz bands

Comparing the radio performance between current cmWave or low mmWave wireless communication systems and future 6G systems operating at frequencies between 100 and 300 GHz require the knowledge of the propagation channel from
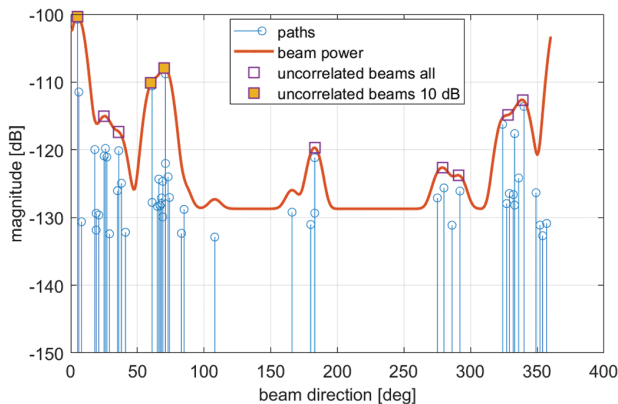
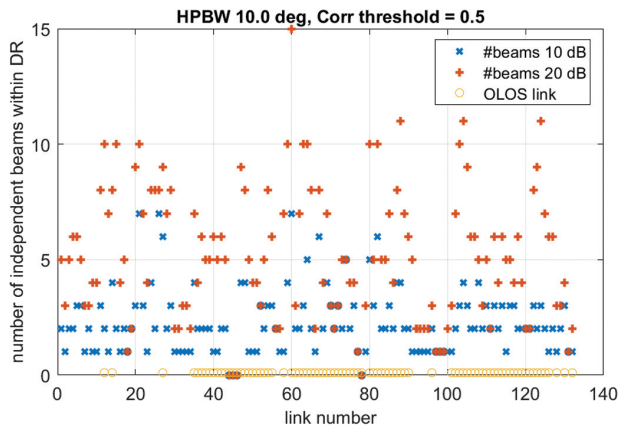**Figure 3.43.** Measured path powers, the beam power, and independent beam azimuth directions in an example link.



**Figure 3.44.** Number of independent beams in 132 indoor links.
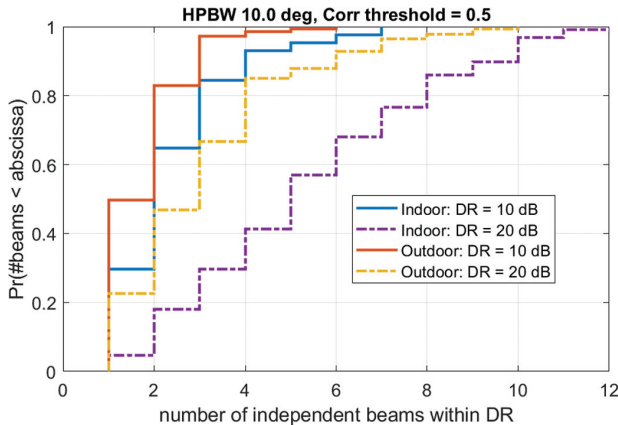


**Figure 3.45.** CDF of number of independent beams in 132 indoor and 157 outdoor links using either 10 or 20 dB dynamic range.

a few GHz to 300 GHz. The propagation channel frequency dependency may be analysed by a pure theoretical approach in free space but needs measurement in more complex environments. Performing wideband channel measurements in real environments over such a bandwidth of several hundred of GHz is technically and organizationally challenging; therefore, a more practical approach is to focus on the material characterization. Knowing the transmission and reflection losses, frequency dependency of usual building material is the first step to assess the potential differences between a propagation channel below 100 GHz and above 100 GHz. A propagation measurement campaign was performed using a vector network analyser (VNA) and frequency extenders to measure continuously from 2 to 260 GHz the reflection (R) and transmission (T) losses of common building material slabs at normal incidence. Figure 3.46 shows results from three typical materials compared with an International Telecommunication Union (ITU)-like model that keeps the ITU theoretical framework for a slab [110] but proposes new parameters fitted by measurement.

Homogeneous and flat surface materials, such as the plexiglass, respect the ITU-fitted model. The observed fading is due to the interference between the direct reflected (or transmitted) path and the multi-paths created inside the material by multiple reflection/transmission on the two air-material interfaces. The
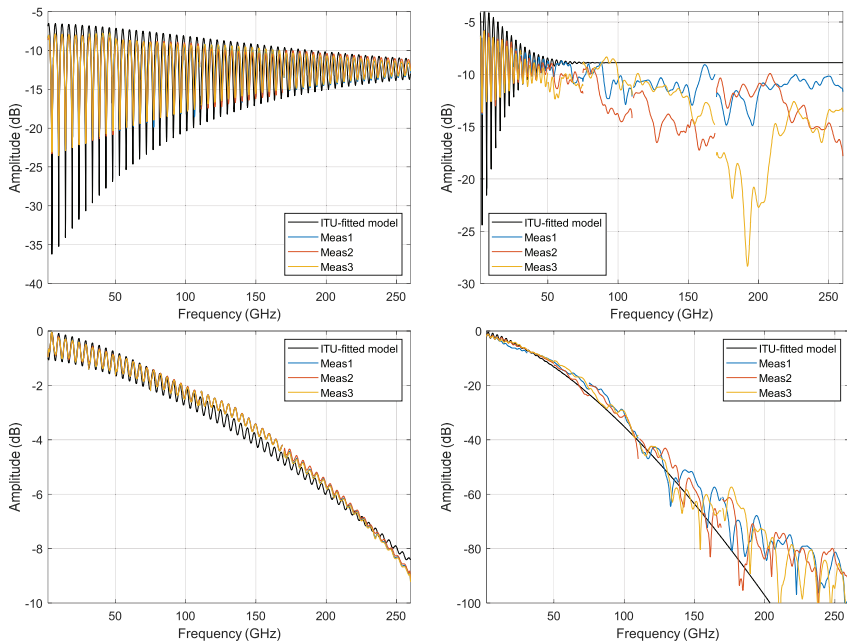


**Figure 3.46.** Material reflection (line a) and transmission gain (line b) compared with the ITU-fitted model fitted by measurement. The measurement is performed at three different points separated by 10 cm on the slab.

fading decreases with the frequency as multi-paths travelling inside the material are attenuated but the R loss average value is constant indicating a constant permittivity. Mortar represents non-homogeneous material with a rough surface. Above 100 GHz, R losses are impacted by the scattering and may be higher than 10 dB compared to a similar material with flat surface. T losses can be still calculated from the conductivity $\sigma$, expressed as $\sigma = cf^d$, $f$ being the frequency and $c$ and $d$ being two parameters. But above 100 GHz, there are strong variations of up to 10 dB around the fitted model. For composite materials such as chipboard or glass wool, R losses can decrease with the frequency increase.

The ITU model is defined up to 100 GHz and needs to be improved for frequencies above 100 GHz from a propagation point of view, but could be used as it is in many simulations related to 6G sub-THz scenarios. Reflection loss errors due to rough surfaces may not be a concern in an office environment, shopping mall, airport, etc., as most of the materials are quite smooth. Transmission loss errors due to the material inhomogeneity may not be a concern as they are related to high transmission losses at high frequencies. Simulating a transmission loss of 20 dBs instead of 30 dBs may not significantly impact system simulation if we consider that a transmission loss higher than 20 dBs corresponds to a blockage. Dedicated multi-frequency measurement at cmWave and sub-THz frequencies in complex environments is required to check these assumptions.

## 3.7   Summary and Outlook

The current chapter provides an overview of selected RAN technologies with a high potential for future 6G networks. It is argued that for the limitless connectivity requirement of 6G, D-MIMO can provide expected macro-diversity (to exploit maximum diversity gain), design flexibility, and interference management, and the related challenges and opportunities are discussed. The main challenge for large-scale D-MIMO roll-out is arguably the cost of installation, as it requires fast and high-speed fronthaul connections. In addition, problems such as beam management aspects, practical approaches to non-coherent operation in higher bands, and transport solutions, e.g., wired/wireless, optical/electrical, and analogue/digital, satisfying the requirements need to be addressed. An optimized wired fronthaul can be realized through radio-stripe/RadioWeaves, while wireless fronthaul/backhaul can be obtained efficiently by IAB. The disaggregation is vital for creating scalable versions of D-MIMO architectures, and thus ORAN support can be viewed as an attractive feature facilitating the progress of the technology. Various ORAN-based deployment scenarios can be considered based on UE clustering and the engagement of O-DUs and O-RUs in serving the clusters. The definition

of relevant deployment scenarios and new interfaces, such as inter-DU interface, is required.

RIS is a key technology for 6G due to its low-cost solution for controlling the propagation channel in favour of the communication link. It shows great potential in several use cases, for not only cellular scenarios, but in UAV and satellite system, and also in IoT scenarios. An architecture and associated three logical entities are introduced in this chapter to facilitate the automated controlling of the RIS, as a main requirement towards its widespread application. Multi-access connectivity is an important and necessary feature for beyond 5G and 6G technologies, for instance in private network scenarios. The 5G CN supports integration of non-3GPP access to provide secure connection for the UE accessing over a non-3GPP access network. Improvements on such framework could be envisioned for more flexible operation using an ORAN-based architecture to benefit the link's throughput, latency, and reliability. The large bandwidth available at mmWave and sub-THz bands introduces the opportunity to overcome the challenge of high-data rate delivery for some demanding 6G use cases. Hardware implementation and modelling the non-linear behaviour of the RF components are a challenge. The modelling of the system is essential for the computation of the link budget and the design of the waveform to mitigate the non-linearity in the system, as well as to estimate the performance of the system.

## References

[1] J.G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A. Lozano, A.C. Soong, J.C. Zhang, "What will 5G be?" In *IEEE J. Sel. Area. Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[2] X. Lin, "An Overview of 5G Advanced Evolution in 3GPP Release 18," In *IEEE Communications Standard Magazine*, vol. 6, no. 3, September 2022.

[3] RP-213468, "Summary for RAN Rel-18 package," 3GPP RAN#94-e, December 2021. Accessed: April 6, 2023. [Online]. Available: http://www.3gpp.org/ftp/tsg_ran/TSG_RAN/TSGR_94e/Docs/RP-213469.zip.

[4] 5GPPP, "Beyond 5G/6G KPIs and Target Values," White Paper, June 2022. Accessed: April 6, 2023. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2022/06/white_paper_b5g-6g-kpis-camera-ready.pdf.

[5] Hexa-X, "D1.2 – Expanded 6G vision, use cases and societal values", Dec. 2021, Accessed April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5dc8b611b&appId=PPGMS.

[6] S. Chen, J. Zhang, J. Zhang, E. Bjornson, B. Ai, "A survey on user-centric cell-free massive MIMO systems," Digital Communications and Networks, 2022.

[7] S. Tripathi, N. V. Sabu, A. K. Gupta, H. S. Dhillon, "Millimeter-wave and Terahertz Spectrum for 6G Wireless," arXiv:2102.10267v1, February 2021. Accessed: April 6, 2023. [Online]. Available: https://arxiv.org/pdf/2102.102 67.pdf.

[8] J. ITU, "Provisional final acts," in World Radiocommunication Conference 2019. ITU Publications, 2019.

[9] Ericsson, "Integrated access and backhaul – a new type of wireless backhaul in 5G," Ericsson Technology Review, June 23, 2020. Accessed: April 6, 2023. [Online]. Available: https://www.ericsson.com/en/reports-and-pa pers/ericsson-technology-review/articles/introducing-integrated-access-and -backhaul.

[10] A. Wolf, P. Schulz, M. Dörpinghaus, J. C. S. Santos Filho, and G. Fettweis, "How reliable and capable is multi-connectivity?," In *IEEE Trans. Commun.* 2018, 67, 1506–1520.

[11] T. Sylla, L. Mendiboure, S. Maaloul, H. Aniss, M. Chalouf, S. Delbruel, "Multi-Connectivity for 5G Networks and Beyond: A Survey," In *MPDI Sensors*, October 2022.

[12] M. Di Renzo, et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," In *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[13] Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable Intelligent Surfaces: Principles and Opportunities," In *IEEE Communications Surveys and Tutorials*," vol. 23, no. 3, pp. 1546–1577, 2021.

[14] N. Rajatheva, et al., "White paper on broadband connectivity in 6G", arXiv preprint, April 2020. Accessed: April 6, 2023. [Online]. Available: https: //arxiv.org/pdf/2004.14247.pdf.

[15] O. Tervo, T. Levanen, K. Pajukoski, J. Hulkkonen, P. Wainio, and M. Valkama, "5G New Radio Evolution Towards Sub-THz Communications," In *2020 2nd 6G Wireless Summit (6G SUMMIT)*, pp. 1–6, 2020. doi: 10.1109/6GSUMMIT49458.2020.9083807.

[16] E. Björnson and L. Sanguinetti, "Cell-Free versus Cellular Massive MIMO: What Processing is Needed for Cell-Free to Win?," In *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, 2019.

[17] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO Versus Small Cells," In *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[18] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of BS antennas," In *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[19] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?", In. *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[20] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," In *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, 2014.

[21] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," In *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, 2018.

[22] D. Gesbert, S. Hanly, H. Huang, S. S. Shamai, O. Simeone, W. and Yu, "Multi-cell MIMO cooperative networks: A new look at interference," In *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, 2010.

[23] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, K. and Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: Deployment scenarios and operational challenges," In *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, 2012.

[24] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO Communications," In *EURASIP J. Wireless Commun. and Networking*, vol. 2019, no. 1, pp. 197–209, 2019a.

[25] *Future Technology Trends of Terrestrial Systems Towards 2030 and Beyond*, Report ITU-R M.251-0, November, 2022.

[26] G. Wikström, P. Persson, S. Parkvall, G. Mildh, E. Dahlman, B. Balakrishnan, P. Öhlen, E. Trojer, G. Rune, J. Arkko, Z. Turányi, D. Roeland, B. Sahlin, W. John, J. Halén, and H. Björkegren, "6G – Connecting a cyber-physical world", Ericsson White Paper, Feb. 2022. Accessed: April 6, 2023. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/a-research-outlook-towards-6g.

[27] *NR; Multi-connectivity; Overall description; Stage-2*, Technical Specification (TS) 37.340, v17.3.0, 3GPP, Jan. 2023.

[28] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," In *IEEE Access*, vol. 7, pp. 99878–99888, 2019.

[29] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5G cellular networks: Resource allocation and comparison with the cell-free massive MIMO approach," In *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1250–1264, 2019.

[30] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," In *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, 2017.

[31] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," In *IEEE Proceedings Communications*, vol. 153, no. 4, pp. 548–555, 2006.

[32] X. Hong, Y. Jie, C.-X. Wang, J. Shi, and X. Ge, "Energy-spectral efficiency trade-off in virtual MIMO cellular systems," In *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2128–2140, Oct. 2013.

[33] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective Multiple Antenna Technologies for Beyond 5G," In *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.

[34] REINDEER "D2.1 Initial assessment of architectures and hardware resources for a RadioWeaves infrastructure," 2022. Accessed: April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e7bcfeeb&appId=PPGMS.

[35] E. Emfors, "Radio Stripes: re-thinking mobile networks," 2019. Accessed: April 6, 2023. [Online]. Available: https://www.ericsson.com/en/blog/2019/2/radio-stripes.

[36] L. Van der Perre, E. G. Larsson, F. Tufvesson, L. De Strycker, E. Björnson, and O. Edfors, "RadioWeaves for efficient connectivity: analysis and impact of constraints in actual deployments," Asilomar Conf., November, 2019.

[37] *NG-RAN; Architecture description*, Technical Specification, TS 38.401 v16.0.0, 3GPP, Jan. 2020,

[38] *ORAN Architecture Description*, Technical Specification (TS) ORAN.WG1.ORAN-Architecture-Description-v04.00, March 2021.

[39] G. Interdonato, P. Frenger, and E.G. Larsson, "Scalability Aspects of Cell-Free Massive MIMO," In *ICC 2019 IEEE International Conference on Communications (ICC)*, vol. 68, no. 7, pp. 1–6, 2019.

[40] V. Ranjbar, A. Girycki, M. A. Rahman, S. Pollin, M. Moonen, and E. Vinogradov, "Cell-Free mMIMO Support in the ORAN Architecture: A PHY Layer Perspective for 5G and Beyond Networks," In *IEEE Communications Standards Magazine*, vol. 6, no. 1, pp. 28–34, March 2022, doi: 10.1109/MCOMSTD.0001.2100067.

[41] *Draft Standard for Ethernet – Amendment: Physical Layer and Management Parameters for DTE Power via MDI over 4-Pair*, IEEE P802.3bt/D1.5, 30 Nov. 2015.

[42] U. K. Ganesan, E. Björnson, and E. G. Larsson, "RadioWeaves for extreme spatial multiplexing in indoor environments," In *2020 54th Asilomar Conference on Signals, Systems, and Computers*, 2020, pp. 1007–1011.

[43] REINDEER "D3.1 Analytical Performance Metrics and Physical-Layer Solutions," 2022. Accessed: April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e7c62ba7&appId=PPGMS.

[44] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, "A Survey on Hybrid Beamforming Techniques in 5G: Architecture and System Model Perspectives," In *IEEE Comm. Surveys & Tutorials*, vol. 20, no. 4, 2018.

[45] *Study on channel model for frequency spectrum above 6 GHz*, Technical Report (TR) 38.900, v14.0.0, 3GPP July 2016.

[46] W. B. Abbas, F. Gomez-Cuba, and M. Zorzi, "Millimeter wave receiver efficiency: A comprehensive comparison of beamforming schemes with low resolution ADCs," In *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, pp. 8131–8146, Dec. 2017.

[47] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-Efficient Hybrid Analog and Digital Precoding for MmWave MIMO Systems with Large Antenna Arrays," In *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[48] C. Fang, B. Makki, J. Li, and T. Svensson, "Hybrid Precoding in Cooperative Millimeter Wave Networks," In *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5373–5388, Aug. 2021.

[49] M. Polese, M. Giordani, T. Zugno, A. Roy, S. Goyal, D. Castor, and M. Zorzi, "Integrated access and backhaul in 5g mmwave networks: Potential and challenges," In *IEEE Communications Magazine*, vol. 58, no. 3, pp. 62–68, 2020.

[50] Techplayon, "5G self-backhaul-integrated access and backhaul," 2019. Accessed: April 6, 2023. [Online]. Available: https://www.techplayon.com/5g-self-backhaul-integrated-access-and-backhaul.

[51] M. N. Islam, S. Subramanian, and A. Sampath, "Integrated access backhaul in millimeter wave networks," In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2017.

[52] A. Ghosh, A. Maeder, M. Baker, D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," In *IEEE Access* vol. 7,

pp. 127639–127651, 2019. Accessed: April 6, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2019.2939938.

[53] O. Teyeb, A. Muhammad, G. Mildh, E. Dahlman, F. Barac, and B. Makki, "Integrated access backhauled networks," In *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, pp. 1–5, 2019.

[54] M. Polese, M. Giordani, A. Roy, D. Castor, and M. Zorzi, "Distributed path selection strategies for integrated access and backhaul at mmwaves," In *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, 2018.

[55] 6G BRAINS, "D4.1 Design and Description of the Intelligent IAB and RmUE/mUE and human-centric control interfaces over Dynamic Ultra-dense D2D Cell Free Network," 2021. Accessed: April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e63588d2&appId=PPGMS.

[56] C. Saha, M. Afshang, and H. S. Dhillon, "Integrated mmWave access and backhaul in 5G: Bandwidth partitioning and downlink analysis," In *Proc. IEEE Int. Conf. Commun. (ICC)*, pp. 1–6, May 2018.

[57] S. M. Azimi-Abarghouyi, B. Makki, M. Haenggi, M. Nasiri-Kenari, and T. Svensson, "Coverage analysis of finite cellular networks: A stochastic geometry approach," In *Proc. Iran Workshop Commun. Inf. Theory (IWCIT)*, pp. 1–5, Tehran, Iran, Apr. 2018.

[58] 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," In *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1374–1389, 2020, doi: 10.1109/OJCOMS.2020.3022529.

[59] C. Madapatha, B. Makki, A. Muhammad, E. Dahlman, M. -S. Alouini, and T. Svensson, "On Topology Optimization and Routing in Integrated Access and Backhaul Networks: A Genetic Algorithm-Based Approach," In *IEEE Open Journal of the Communications Society*, vol. 2, pp. 2273–2291, 2021, doi: 10.1109/OJCOMS.2021.3114669.

[60] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable intelligent surfaces: A signal processing perspective with wireless applications," In *arXiv preprint,* 2021, arXiv:2102.00742.

[61] E. C. Strinati, G. C. Alexandropoulos, Vincenzo Sciancalepore, M. Renzo, H. Wymeersch, D. P. Huy, M. Crozzoli, R. D'Errico, E. Carvalho, P. Popovski, P. Lorenzo, L. Bastianelli, Mathieu Belouar, J. Mascolo, G. Gradoni, S. Phang, G. Lerosey, and B. Denis, "Wireless environment as a service enabled by reconfigurable intelligent surfaces: The RISE-6G perspective," In *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2021.

[62] E. C. Strinati, G. C. Alexandropoulos, H. Wymeersch, B. Denis, V. Sciancalepore, R. D'Errico, A. Clemente, D. T. Phan-Huy, E. De Carvalho, and P. Popovski, "Reconfigurable, Intelligent, and Sustainable Wireless Environments for 6G Smart Connectivity," in IEEE Communications Magazine, vol. 59, no. 10, pp. 99–105, October 2021, doi: 10.1109/MCOM.001.2100070.

[63] Y. Cao and T. Lv, "Intelligent reflecting surface enhanced resilient design for MEC offloading over millimeter wave links," 2019. arXiv:1912.06361.

[64] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," In *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[65] S. Li, B. Duo, X. Yuan, Y. C. Liang, and M. Di Renzo, "Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming," In *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 716–720, May 2020.

[66] D. Ma, M. Ding, and M. Hassan, "Enhancing cellular communications for UAVs via intelligent reflective surface," In *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 1–6, 2020.

[67] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Intelligent reflecting surface enhanced indoor robot path planning: A radio map-based approach," In *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, July 4, 2021, doi: 10.1109/TWC.2021.3062089.

[68] RISE-6G "D2.5: RISE network architectures and deployment strategies analysis: first results," 2022. Accessed: April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5ef286723&appId=PPGMS.

[69] A. Albanese, F. Devoti, V. Sciancalepore, M. Di Renzo, and X. Costa-Pérez, "MARISA: A self-configuring metasurfaces absorption and reflection solution towards 6G," In *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*, pp. 250–259, May 2022.

[70] I. Alamzadeh, G. C. Alexandropoulos, N. Shlezinger, and M. F. Imani, "A reconfigurable intelligent surface with integrated sensing capability," In. *Sci. Rep.* vol. 11, pp. 20737, 2021. Accessed: April 6, 2023. [Online]. Available: https://doi.org/10.1038/s41598-021-99722-x.

[71] G. Stratidakis, S. Droulias, and A. Alexiou, "An analytical framework for reconfigurable intelligent surfaces placement in a mobile user environment," In *Proceedings of the 19th ACM Conference on Embedded Networked Sensor Systems, ser. SenSys '21*. New York, NY, USA: Association for Computing

Machinery, pp. 623–627, 2021. Accessed: April 6, 2023. [Online]. Available: https://doi.org/10.1145/3485730.3494038.

[72] G. Stratidakis, S. Droulias, and A. Alexiou, "Optimal position and orientation study of Reconfigurable Intelligent Surfaces in a mobile user environment," In *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 10, 2022, doi: 10.1109/TAP.2022.3208036.

[73] A.-A. A. Boulogeorgos, N. D. Chatzidiamantis, H. G. Sandalidis, A. Alexiou and M. D. Renzo, "Cascaded Composite Turbulence and Misalignment: Statistical Characterization and Applications to Reconfigurable Intelligent Surface-Empowered Wireless Systems," In *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3821–3836, April 2022, doi: 10.1109/TVT.2021.3140084.

[74] A.-A. A. Boulogeorgos, N. D. Chatzidiamantis, H. G. Sandalidis, A. Alexiou, and M. D. Renzo, "Performance Analysis of Multi-Reconfigurable Intelligent Surface-Empowered THz Wireless Systems," In *IEEE International Conference on Communications (ICC)*, 2022.

[75] C. Liaskos, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "Using any surface to realize a new paradigm for wireless communications," In *Communications of the ACM*, vol. 61, no. 11, pp. 30–33, Oct. 2018.

[76] A. -A. A. Boulogeorgos, A. Alexiou, and M. D. Renzo, "Outage performance analysis of RIS-assisted UAV wireless systems under disorientation and misalignment," In *IEEE Transactions on Vehicular Technology*, 2022, doi: 10.1109/TVT.2022.3187050.

[77] 3GPP, "Study on Access Traffic Steering, Switch and Splitting Support in the 5G System Architecture," 3GPP Technical Report, TR 23.793, 2018.

[78] 5G-CLARITY "D3.2 Design Refinements and Initial Evaluation of the Coexistence, Multi-Connectivity, Resource Management and Positioning Frameworks," 2021. Accessed: April 6, 2023. [Online]. Available https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e8fc7cff&appId=PPGMS.

[79] A. Ford, C. Raiciu, M. J. Handley, and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses," RFC 6824, Jan. 2013. Accessed: April 6, 2023. [Online]. Available: https://www.rfc-editor.org/info/rfc6824.

[80] C. Paasch, F. Duchêne, and G. Detal, "Multipath TCP – Linux Kernel implementation," 2023. Accessed: April 6, 2023. [Online]. Available: https://multipath-tcp.org/.

[81] 5G-CLARITY "D3.3 Complete Design and Final Evaluation of the Coexistence, Multi-Connectivity, Resource Management, and Positioning Frameworks," 2022. Accessed: April 6, 2023. [Online]. Available: https://ec.europ

a.eu/research/participants/documents/downloadPublic?documentIds=080166e5f3094463&appId=PPGMS.

[82] Virtualized multi-connectivity testbed, 5G-CLARITY Project. Accessed: April 6, 2023. [Online]. Available: https://github.com/jorgenavarroortiz/multitechnology_testbed_v0.

[83] A. Bourdoux, A. N. Barreto, B. van Liempd, C. de Lima, D. Dardari, D. Belot, E. S. Lohan, G. Seco-Granados, H. Sarieddeen, H. Wymeersch, J. Suutala, J. Saloranta, M. Guillaud, M. Isomursu, M. Valkama, M. R. K. Aziz, R. Berkvens, T. Sanguanpuak, T. Svensson, and Y. Miao, "6G white paper on localization and sensing." In *arXiv preprint* arXiv:2006.01779, 2020.

[84] M. E. Leinonen, M. Jokinen, N. Tervo, O. Kursu, and A. Pärssinen, "System EVM Characterization and Coverage Area Estimation of 5G Directive mmW Links," In *IEEE Trans. on Microw. Theory and Techn.*, vol. 67, no. 12, pp. 5282–5295, Dec. 2019.

[85] Hexa-X, "D2.1 – Towards Tbps Communications in 6G: Use cases and Gap Analysis", Dec. 2021, Accessed April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5df2529fd&appId=PPGMS.

[86] Hexa-X, "D2.2 – Initial radio models and analysis towards ultra-high data rate links in 6G", Dec. 2021, Accessed April 6, 2023. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5e648d315&appId=PPGMS.

[87] A. Karakuzulu, A. Malignaggi, and D. Kissinger, "Low Insertion Loss D-band SPDT Switches Using Reverse and Forward Saturated SiGe HBTs," 2019 IEEE Radio and Wireless Symposium (RWS), 2019, pp. 1–3, doi: 10.1109/RWS.2019.8714362.

[88] X. Zhao, O. Glubokov, J. Campion, A. Gomez-Torrent, Aleksandr Krivovitca, U. Shah, and J. Oberhammer, "Silicon Micromachined D-Band Diplexer Using Releasable Filling Structure Technique," In *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 8, pp. 3448–3460, Aug. 2020, doi: 10.1109/TMTT.2020.3004585.

[89] H. Wang, et al., "Power Amplifiers Performance Survey 2000-Present," Georgia Tech. Accessed: April 6, 2023. [Online]. Available: https://gems.ece.gatech.edu/PA_survey.html.

[90] L. Belostotski, S. Jagtap, "Down With Noise: An Introduction to a Low-Noise Amplifier Survey," In *IEEE Solid-State Circuits Magazine*, vol. 12, no. 2, pp. 23–29, Spring 2020, doi: 10.1109/MSSC.2020.2987505.

[91] E. Dalman, S. Parkval, J. Skold, 5G NR, *The Next Generation Wireless Access Technology*, 2nd ed., Academic Press 2020, Chapter 26.

[92] B. Murmann, "ADC Survey," Accessed April 6, 2023. [Online]. Available: https://github.com/bmurmann/ADC-survey.

[93] P. Skrimponis, S. Dutta, M. Mezzavilla, S. Rangan, S. H. Mirfarshbafan, C. Studer, J. Buckwalter, and M. Rodwell, "Power Consumption Analysis for Mobile MmWave and Sub-THz Receivers," In *2020 2nd 6G Wireless Summit (6G SUMMIT)*, pp. 1–5, 2020. doi: 10.1109/6GSUMMIT49458.2020.9083793.

[94] A. Pärssinen, M. Alouini, M. Berg, T. Kuerner, P. Kyösti, M. E. Leinonen, M. Matinmikko-Blue, E. McCune, U. Pfeiffer, and P. Wambacq, (Eds.), "White Paper on RF Enabling 6G – Opportunities and Challenges from Technology to Spectrum," [White paper]. (6G Research Visions, No. 13). University of Oulu, 2020. Accessed: April 6, 2023. [Online]: Available: http://urn.fi/urn:isbn:9789526228419.

[95] P. Rodriguez-Vazquez, J. Grzyb, B. Heinemann, and U. R. Pfeiffer, "A QPSK 110-Gb/s Polarization-Diversity MIMO Wireless Link With a 220–255 GHz Tunable LO in a SiGe HBT Technology," In *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 9, pp. 3834–3851, Sept. 2020.

[96] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Pearson Education Ltd., 3rd Ed., London, UK, p. 1056, 2013.

[97] D. Schreurs, M. O'Droma, A. A. Goacher, M. Gadringer et al., "Power Amplifier Behavioral Modeling," Cambridge University Press, p. 263, 2009.

[98] L. Belostotski and S. Jagtap, "Low-noise amplifier (LNA) performance survey," Univ. of Calgary, Canada, Jan. 2020. Accessed: April 6, 2023. [Online] Available: https://profiles.ucalgary.ca/sites/default/files/2022-08/lna_survey.xlsx.

[99] P. Rodríguez-Vázquez, M. E. Leinonen, J. Grzyb, N. Tervo, A. Parssinen, and U. R. Pfeiffer, "Signal-processing Challenges in Leveraging 100 Gb/s Wireless THz," In *2020 2nd 6G Wireless Summit (6G SUMMIT)*, pp. 1–5, 2020.

[100] *Study on supporting NR from 52.6 GHz to 71 GHz*, Technical Report (TR) 38.303, v17.0.0, 3GPP, March 2021.

[101] G. Fettweis, M. Dörpinghaus, S. Bender, L. Landau, P. Neuhaus, and M. Schlüter, "Zero Crossing Modulation for Communication with Temporally Oversampled 1-Bit Quantization," In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pp. 207–214, 2019. doi: 10.1109/IEEECONF44664.2019.9048794.

[102] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier FDMA for uplink wireless transmission," In *IEEE Vehicular Technology Magazine*, vol. 1, no. 3, pp. 30–38, Sept. 2006, doi: 10.1109/MVT.2006.307304.

[103] M. P. Wylie-Green, E. Perrins, T. Svensson, "Introduction to CPM-SC-FDMA – A Novel Multiple-Access Power-Efficient Transmission Scheme." In *IEEE Transactions on Communications*, vol. 7, pp. 1904–1915, 2011.

[104] M. F. De Guzman, P. Koivumäki, and K. Haneda, "Double-directional Multipath Data at 140 GHz Derived from Measurement-based Ray-launcher," in Proc. 2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring), 2022, pp. 1–6.

[105] A. F. Molisch, M. Steinbauer, M. Toeltsch, E. Bonek, and R. S. Thoma, "Capacity of MIMO systems based on measured wireless channels," In *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 561–569, April 2002.

[106] M. F. De Guzman and K. Haneda, "Double-directional multipath data at 140 GHz," available: 10.5281/zenodo.6653867.

[107] S. L. H. Nguyen, J. Järveläinen, A. Karttunen, K. Haneda, and J. Putkonen, "Comparing radio propagation channels between 28 and 140 GHz bands in a shopping mall," In *Proc. 12th European Conf. Ant. Prop. (EuCAP 2018)*, London, UK, pp. 1–5, Apr. 2018.

[108] P Kyösti, M. F. De Guzman, K Haneda, N Tervo, A Pärssinen, "How many beams does sub-THz channel support?," In *IEEE Antennas and Wireless Propagation Letters*, vol. 21, no. 1, pp. 74–78, 2021.

[109] *Study on channel model for frequencies from 0.5 to 100 GHz*, Technical Report (TR) 38.901 v14.1.1, 3GPP, Jul. 2017.

[110] *Effects of building materials and structures on radio wave propagation above about 100 MHz*, Recommendation ITU-R P.2040-1, ITU-R, Jul. 2015.

[111] A. U. Makarfi, K. M. Rabie, O. Kaiwartya, O. S. Badarneh, X. Li, and R. Kharel, "Reconfigurable intelligent surface enabled IoT networks in generalized fading channels," In *Proc. IEEE Int. Commun. Conf. (ICC)*, pp. 1–6, 2020.

[112] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," In *IEEE Trans. Green Commun. and Netw.*, vol. 2, no. 1, pp. 25–39, 2018.

[113] J. Liu, M. Sheng, L. Liu, and J. Li, "Network Densification in 5G: From the Short-Range Communications Perspective," In *IEEE Communications Magazine*, vol. 55, no. 12, pp. 96–102, December 2017.

[114] F. Sohrabi and W. Yu, "Hybrid Digital and Analog Beamforming Design for Large-Scale Antenna Arrays," In *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, April 2016, doi: 10.1109/JSTSP.2016.2520912.