Call: H2020-ICT-2020-2

Project reference: 101015956

Project Name:

A flagship for B5G/6G vision and intelligent fabric of technology enablers connecting human, physical, and digital worlds

Hexa-X

Deliverable D2.3
Radio models and enabling techniques towards ultra-high data rate links and capacity in 6G

Date of delivery:	31/03/2023	Version:	1.0
Start date of project:	01/01/2021	Duration:	30 months

Document properties:

Document Number:	D2.3
Document Title:	Radio models and enabling techniques towards ultra-high data rate links and capacity in 6G
Editor(s):	Ahmad Nimr (TUD)
Authors:	Ahmad Nimr (TUD), Marko E. Leinonen (OUL), Muris Sarajlic (EAB), Tommy Svensson (CHA), Katsuyuki Haneda (AAU), Andreas Wolfgang (QRT), Han Yu (CHA), Charitha Madapatha (CHA), Hao Guo (CHA), Pål Frenger (EAB), Rafael Puerta (EAB), Behrooz Makki (EAB), Yinggang Li (EAB), Omer Haliloglu (EBY), Kilian Roth (INT), Leonardo Gomes Baltar (INT), Hardy Halbauer (NOG), Le Hang Nguyen (NOG), Mohamed Abdelbasset Aliouane (ORA), Jean-Marc Conrat (ORA), Pekka Kyösti (OUL), Nuutti Tervo (OUL), Aarno Pärssinen (OUL), Xinxin Yang (QRT), Meik Dörpinghaus (TUD), Stephan Zeitz (TUD), Roberto Bomfin (TUD)
Contractual Date of Delivery:	31/03/2023
Dissemination level:	PU[1]/
Status:	Final
Version:	1.0
File Name:	Hexa-X D2.3

Revision History

Revision	Date	Issued by	Description
v0.1	12.04.2022	Ahmad Nimr	Initial Version
v0.2	09.09.2022	Ahmad Nimr	Section 2 changes accepted
V0.3	01.09.2023	Ahmad Nimr	Clean up draft for internal review
V0.4	16.02.2023	Ahmad Nimr	Internal review during F2F
V0.5	20.02.2023	Ahmad Nimr	For external review
V1.0	31.03.2023	Ahmad Nimr	Release version

[1] CO = Confidential, only members of the consortium (including the Commission Services)

PU = Public

Abstract

This report presents the findings of studies and performance analysis on 6G radio system for achieving ultra-high data rate links and capacity in the context of Hexa-X work package 2: "Novel radio access technologies towards 6G". The report begins by presenting an overview of use cases relevant to sub-THz communication, defining representative scenarios, technical requirements, and performance metrics. Then, technical studies follow in the areas of channel modelling, radio architecture including transceiver design and distributed MIMO, in addition to signal processing and transmission schemes. In particular, radio transceiver design approaches are explored considering hardware non-idealities, channel properties, and waveform impacts. The report introduces description and analysis based on measurement for different hardware components and channel model using measurement data, which are used in the design and analysis of the physical layer modulation and waveforms. Furthermore, D-MIMO architecture options and corresponding transmission and processing schemes are investigated with performance analysis focusing on coverage and throughput. Finally, system-level simulation is conducted to illustrate the impact of different deployment options and design parameters on the performance.

Keywords

100 – 300 GHz, hardware model, beamforming, channel measurements, distributed MIMO, physical layer, radio model, simulation model

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101015956.

Executive Summary

This report is the third deliverable of Hexa-X project work package 2 (WP2): "Novel radio access technologies towards 6G". It focuses on 6G radio system design aspects and enabling techniques for fulfilling the requirements of emerging use cases, especially, link data rate of 100 Gbps. These include radio frequency (RF) transceiver architectures and hardware (HW) components models for frequency above 100 GHz, channel measurements and modelling, deployment scenarios with cellular-based and distributed MIMO (D-MIMO) approaches, and signal processing topics on waveform, beam management, and D-MIMO transmission schemes. Moreover, performance analysis is presented with different objectives, such as link-level evaluation of modulation under HW nonidealities, and system-level analysis to assess the performance of different design and deployment options.

The report starts with an overview of a set of use cases enabled by sub-THz communication in the band (100 GHz-300 GHz) introduced in deliverable [HEX21-D21] with key performance indicators (KPIs) requirements for data rate and latency. Based on range requirements, three scenarios (mid-range, short range, and very short range) are defined with detailed specifications including the positioning requirements from WP3 report [HEX21-D31]. For each scenario, several radio options are foreseen depending on the operating band and device class (user equipment (UE), or access point (AP)). To narrow the design space, system performance metrics are used to evaluate the design and deployment scenarios, including coverage, energy performance, link reliability and latency, beam failure rate, and positioning metrics, for the. Moreover, design considers the impacts of radio channel, including path loss, fading, and opportunities for spatial beams. This is based on the results of channel models for material interaction in the range (2-260 GHz), and a using the stored channel model data at 140 GHz.

The report then presents technical guidelines for RF transceiver design to achieve link data rate at a particular range with detailed discussions and evaluation examples for each step. Various parameters and degrees of freedom are discussed, including bandwidth, signal-to-noise ratio (SNR), analogue-to-digital convertors (ADC), HW nonidealities, and waveform impact on the power amplifier (PA) backoff. Moreover, it discusses RF implementation aspects and strategies to optimize the system such as channelization and carrier aggregation (CA). Based on the analysis, it is concluded that sub-array-based RF transceiver is a practically feasible architecture for above 100 GHz, which implies analogue beam steering per subarray and digital precoding over multiple RF chains. Different configurations for the RF chains are possible; e.g., spatial multiplexing, or implementation of ultra-wideband waveforms following CA approach. However, conventional initial beam access is impractical due to the large number of beams to be probed, and it is necessary to consider side information for beam management.

The document describes several hardware models for local oscillator (LO) phase noise and PA, and their accuracy is justified based on measurements and circuit analysis. These models are considered in the signal processing analysis to design waveforms and modulation schemes that mitigate HW non-idealities, or to develop solutions for estimation and compensation. Accordingly, performance analysis is presented for one or more HW impairments, with special focus on DFTS-OFDM and single carrier (SC), as they show tolerance to PA nonidealities compared to OFDM [HEX21-D22], and zero-crossing modulation (ZXM) with 1-bit quantization as a potential energy-efficient modulation scheme candidate.

With respect to other performance metrics, a study on RAN latency to achieve 0.1 ms is conducted, showing that processing latency with high throughput is the most critical factor. This needs to be considered in the design of processing architecture considering I/O speeds, memory access speed, and power consumption. The report also provides several performance analyses in the context of D-MIMO architecture, focusing on spectral efficiency, throughput, and coverage for different transmission schemes and techniques to cope with blockage. Finally, a system-level simulation example is studied to illustrate the impact of radio configuration and deployment options on the performance in terms of area throughput and energy consumption, showing the potential for optimizing radio design in terms of transceiver architecture and deployment scenario to achieve KPIs at lower energy cost.

Table of Contents

List of Figures

List of Tables

List of Acronyms

Term	Description
3D	3 Dimensional
3GPP	3rd Generation Partnership Project
5G	5th Generation mobile communication
5G-PPP	5G Public Private Partnership
6G	6th Generation mobile communication
ABG	Alpha-Beta-Gamma
ACLR	Adjacent Channel Leakage Power
ADC	Analogue-to-Digital Converter
AGC	Automatic Gain Control
AM-AM	Amplitude-to-Amplitude
AM-PM	Amplitude-to-Phase
AoA	Angle-of-Arrival
AoD	Angle-of-Departure
AP	Access Point
APSK	Amplitude Phase Shift Keying
AR	Augmented Reality
ARoF	Analog Radio over Fibre
AWG	Arbitrary Waveform Generator
AWGN	Additive White Gaussian Noise
B5G	Beyond 5G
BB	Baseband
BCRB	Bayesian Cramér-Rao bound.
BER	Bit Error Rate
BFW	Beamforming Weights
BiCMOS	Bipolar CMOS
BLER	Block Error Rate
BM	Beam Management
BS	Base Station
BW	Bandwidth
CA	Carrier Aggregation
CC	Component Carriers
CCDF	Complementary Cumulative Density Function
CI	Close-In
CJT	Coherent Joint Transmission
CMOS	Complementary Metal Oxide Semiconductor
CPE	Common Phase Error

CP-OFDM	Cyclic Prefix OFDM
CPRI	Common Public Radio Interface
CRC	Cyclic Redundancy Check
CSI	Channel State Information
CU	Centralized Unit
DAC	Digital-to-Analogue Converter
DC	Drained Current
DFT	Discrete Fourier Transformation
DFTS-OFDM	DFT Spread OFDM
DL	Downlink
D-MIMO	Distributed MIMO
DSO	Digital Storage Oscilloscope
DSP	Digital Signal Processing
DT	Digital Twin
DU	Distributed Unit
DUT	Device Under Test
DVB-S2	Digital video broadcasting second generation
E2E	End-to-End
EA	Electrical Amplifier
EC	European Commission
ECL	External Cavity Laser
EIRP	Effective Isotropic Radiated Power
EM	Electromagnetic
EU	European Union
EVM	Error Vector Magnitude
FER	Frame Error Rate
FH	Fronthaul
FOM	Figure Of Merit
FR1	Frequency Range 1
FR2	Frequency Range 2
FTN	Faster Than Nyquist
GaAs	Gallium Arsenide
GaN	Gallium Nitride
Gbps	Giga bits per second
GFB	Gypsum Fibre Board
GoB	Grid of Beams
Gsps	Giga symbol per second
GTD	Geometrical Theory of Diffraction
H2020	Horizon 2020

HW	Hardware
HWI	Hardware Impairment
i.i.d.	independent and identically distributed
IAB	Integrated Access and Backhaul
IBKM	Information-Based k-Means
IC	Integrated Circuit
ICI	Inter-Carrier Interference
ICP	Input Compression Point
IF	Intermediate Frequency
IFFT	Inverse Fast Fourier Transform
InP	Indium Phosphide
IQ	In-phase and Quadrature
ISI	Inter Symbol Interference
KPI	Key Performance Indicator
LCA	Life Cycle Assessment
LDPC	Low-Density Parity-Check
LINC	Linear Amplification with Nonlinear Components
LLS	Lower Layer Split
LNA	Low Noise Amplifier
LO	Local Oscillator
LOS	Line-of-Sight
MAC	Medium Access Control
MCL	Maximum Coupling Loss
MDF	Medium Density Fibreboard
MDS	Maximum Distance Separable
MIMO	Multiple Input Multiple Output
MMSE	Minimum Mean Square Error
mmWave	Millimetre-wave
MR	Mixed Reality
MSE	Mean-Squared Error
MT	Mobile Termination
MTC	Machine Type Communication
MZM	Mach-Zehnder Modulator
NCJT	Non-coherent Joint Transmission
NCR	Network-Controlled Repeater
NF	Noise Figure
NLOS	Non-Line-of-Sight
NR	New Radio
NRZI	Non-Return-to-Zero Inverse

OBO	Output Backoff
OFDM	Orthogonal Frequency Division Multiplexing
OOB	out-of-band
OTA	Over-the-Air
PA	Power Amplifier
PAPR	Peak to Average Power Ratio
PFB	Plastic Fibre Board
PLL	Phase Locked Loop
PN	Phase Noise
PSD	Power Spectral Density
PTRS	Phase Tracking Reference Signal
PU	Peer Unit
QAM	Quadrature Amplitude Modulation
RAN	Radio Access Network
RF	Radio Frequency
RFIC	Radio Frequency Integrated Circuit
RIS	Reconfigurable Intelligent Surfaces
RLL	Runlength-Limited
RMCL	Relative Maximum Coupling Loss
RMS	Root Mean Square
RRC	Root Raised Cosine
RSSD	Reduced State Sequence Detection
RTS	Rauch-Tung-Striebel
RU	Radio Unit
Rx	Receiver
RX	Reception
SC	Single Carrier
SC-FDE	Single-Carrier Frequency Domain Equalization
SCS	Subcarrier Spacing
SE	Spectral Efficiency
SER	Symbol Error Rate
SiGe	Silicon germanium
SINR	Signal-to-Interference and Noise-Ratio
SISO	Single Input Single Output
sMWIM	Sequential Maximum Weighted Induced Matching
SNDR	Signal-to-Noise and Distortion-Ratio
SNR	Signal-to-Noise Ratio
SoA	State-of-the-Art
SRS	Sounding Reference Signal

SRT	Spatial Repetition Transmission
SSB	Single-Sideband
SSMF	Standard Single-Mode Fibre
TDD	Time Division Duplex
TM	Test Model
ToA	Time-of-arrival
TRP	Total Radiated Power
TTI	Transmission Time Interval
Tx	Transmitter
TX	Transmission
UE	User Equipment
UL	Uplink
UPA	Uniform Planar Array
UTD	Uniform Theory of Diffraction
UWB	Ultrawideband
VCO	Voltage Controlled Oscillator
VNA	Vector Network Analyser
w.r.t.	with respect to
WDM	Wavelength Division Multiplex
WFH	Wireless Fronthauling.
WP	Work Package
XR	Extended Reality
ZF	Zero Forcing
ZoA	Zenith Angle-of-Arrival
ZoD	Zenith Angle-of- Departure
ZXM	Zero-crossing Modulation

1 Introduction

Hexa-X is one of the 5G-PPP projects under the EU Horizon 2020 framework. It is a flagship project that develops a Beyond 5G (B5G)/6G vision and an intelligent fabric of technology enablers connecting human, physical and digital worlds.

This document is the third deliverable (D2.3) of Work Package 2 (WP2) - "Novel radio access technologies towards 6G". The work in WP2 focuses on radio design aspects for future mobile communication systems, including radio frequency (RF) transceiver implementation and hardware (HW) modelling in sub-THz bands (100 – 300 GHz), distributed multiple-input multiple-output (D-MIMO) architectures, and signal processing and transmission schemes. The research focuses on the following key aspects:

- Radio enablers and technology roadmap
- Radio and antenna implementation aspects, hardware component models and architecture
- Hardware-aware waveform and modulation design
- Hardware-aware beamforming design
- Distributed large MIMO systems for beyond 5G and 6G
- Channel measurement and modelling above 100 GHz

The first deliverable of WP2 (D2.1) [HEX21-D21] presents use cases related to Tbps communications, and identifies gaps in the current and future technologies, which steer the work in WP2. The second deliverable (D2.2) [HEX21-D21] provides initial analysis of radio models and performance analysis at the radio system level, modelling of non-linearities, channel measurement and modelling, and waveform analysis with the modelled radio impairments.

1.1 Objective of the document

This deliverable aims at finalizing the radio models introduced in D2.2 and providing performance results based on link and system-level analysis and simulation for the use cases presented in D2.1. The main objective is to develop radio design methodology and evaluate enabler technologies for achieving ultra-high data rate for links, and high system capacity, considering other performance metrics such as energy efficiency. The first option focuses on wireless technologies operating at the frequency range of 100 to 300 GHz, where design guidelines for RF transceiver is presented considering implementation aspects, HW models, and modulation schemes. The second option is based on D-MIMO at frequencies below 100 GHz, where various architectures and transmission schemes are studied and evaluated.

The results in this report can be used as basis for future works, especially, for further optimization of radio design to meet sustainability targets.

1.2 Structure of the document

The remainder of this document is structured as follows: Section 2 presents an overview of the relevant use cases, and defines the corresponding scenarios based on the link range with detailed requirements. Moreover, it presents an overview of the generic radio architecture and a general design methodology. Furthermore, deployment options using cellular-based and D-MIMO are discussed.

Section 3 provides studies and description of two channel models, the first is material parameters for 2-260 GHz, and the second is a stored channel model at 140 GHz that is based on measured channel responses as a seed to reproduce channels based on a Monte Carlo simulation.

Section 4 focuses on radio architecture and modelling aspects. The first part focuses on system design of RF transceiver for the frequency range (100 – 300 GHz), and provides design guidelines. The second part is dedicated for the description and evaluation of the HW models, in particular, for local oscillator (LO) phase noise, power amplifier (PA). In the third part, D-MIMO architectures are presented.

In Section 5, signal processing topics related to the design of high-throughput 6G systems are presented. First, an overview is given of selected waveform design topics and related digital transceiver design guidelines. Then, guidelines are given on how the beam management techniques need to be adapted to the large number of beams needed in sub-THz systems. Thereafter, a look into a selection of topics in the area of D-MIMO and integrated access and backhaul is provided.

Section 6 is dedicated for system performance analysis, where the impact of deployment scenarios on the power consumption is evaluated. In addition, insights are provided on the influence of the propagation channel on the link quality based on channel measurements in several environments.

The document is concluded in Section 7.

2 Scenarios and technical requirements

As foreseen by Hexa-X [HEX21-D11] and [HEX21-D51], the 6G architecture will be a network of networks that are established for different use cases and deployment scenarios. Among the subnetworks, the sub-THz radio access network (RAN) has the potential to realize use cases with extreme performance requirements in terms of data rate and latency. However, because of the challenges associated with the design of radio hardware at the sub-THz upper Millimetre-wave (mmWave) (100-300 GHz) [HEX21-D21], and tight interdependence of communication range with both propagation environment and applicable use cases, the radio design needs to be optimized according to the expected range and device level power and performance constraints. The radio design at sub-THz frequencies and ultrawide bandwidth needs to consider the complexity of multi-antenna implementations and beamforming, the radio frequency HW impairments, the limitation of the transmit power, and the properties of the propagation channel. In addition, due to the variety of use cases and deployment scenarios, a single radio HW will not be necessarily an ideal solution to serve all the scenarios.

The radio HW design is determined by the required data rate, maximum communication range, expected device mobility/time variability of the channel, expected number of user equipment (UE) to be served and the number of independent spatial directions provided by the propagation environment, which can be exploited to increase the achievable data rate while relaxing requirements on the bandwidth and transmitted power. The design should also consider the sustainability impacts including the power consumption and material life cycle assessment (LCA). A proper design provides the signal-to-noise ratio (SNR) range required by the employed waveforms and transmission schemes. Accordingly, the impact of the waveform needs to be evaluated in advance in the radio analyses and design phase. Once the radio architecture is selected, the waveform and frame design should be designed to fulfil the air interface latency requirement, which is a portion of the end-to-end (E2E) latency budget, and it includes retransmission delay, buffering delay, and processing delay.

2.1 Representative communication scenarios for sub-THz

Several emerging use cases are relevant to sub-THz as they require extremely high data rate and low latency [HEX21-D13], in addition to the stringent sensing requirements as in [HEX21-D31]. Moreover, sub-THz can be considered to provide fixed access to replace optical fibres [HEX21-D21]. This section provides a short summary of the targeted use cases and explain the investigated sub-THz communication scenarios and requirements that concern the radio HW design and air interface.

2.1.1 Relevant use cases

Digital twins for manufacturing: a digital twin (DT) is a virtual replica of a physical asset and provides a digital representation in time and space. The actual status of the physical twin needs to be reflected accurately in real-time. This requires collecting and conveying massive amount of data from different sensors at high data rate. DT can be used in different indoor and outdoor environments, such as ports, mines, and factories. A small cell or sensor network coverage is required for connecting a massive number of different types of devices, including sensors, alarms, fixed machinery, and moving autonomous vehicles. These devices can be connected to a fixed or a mobile gateway with controllable mobility in terms of location and trajectory, e.g., by means of remote operation or predefined settings. The gateway is connected to the access point via a sub-THz link. The required peak data rate is 10-100 Gbit/s with E2E latency in the range of 0.1-100 ms. The positioning accuracy is on the order of 1 cm and latency of 0.1-100 ms.

Fully merged cyber-physical worlds: this use case goes beyond the conventional audio and visual immersive experience to include haptic communication and allows mixed reality of the virtual and physical environment. Among the applications of this use case are holographic communication, telepresence, augmented shopping, guidance and information, mixed reality (MR) entertainment in gaming and tourism, and high-rate local wireless access. The deployment of the network can be indoor

malls and subways, or outdoor streets and parks. While considering multiple users, the deployment option should be high density cells. The involved devices include augmented reality (AR) glasses and tactile gloves that are locally connected to a terminal that is connected to an access point (AP). As the devices are held, or worn, by a human, the mobility is at walking speed and can be stationary. Depending on the application, the data rate ranges from 1 Gbps to 100 Gbps with E2E latency less than 20 ms. The positioning accuracy should be on the order of 1 cm and the latency of less than 100 ms.

Fixed wireless access and fronthaul: In this case, sub-THz can be used to provide directional wireless link between stationary nodes with high data rate (100 Gbps) and low latency (0.1-100 ms) as alternative to optical fibres for fronthaul links for interconnecting small cell APs, connecting remote digital twins, and providing fixed wireless access for residential users. Such link is commonly deployed in outdoor environment between stationary APs, but can also be relevant to indoor. For supporting the sensor infrastructure web, location accuracies around 0.1 - 1 meter with latencies of 10-100 ms are required.

2.1.2 Sub-THz scenarios

The aforementioned use cases are translated to three scenarios characterized by the key performance indicators (KPIs): maximum range, data rate, and E2E latency, as listed in Table 2-1. In addition, the deployment and operating conditions including mobility, type of radio channel, and type of involved devices. Considering the type of devices, the radio HW design can be symmetric (the same architecture at both transmitter (Tx) and receiver (Rx)) or asymmetric (different architectures for Tx and Rx), where the devices constraints need to be considered. In fact, both symmetric and asymmetric design can be investigated in all cases, and for different type of devices. Note that this symmetry is not related to uplink/downlink (UL/DL) requirement because time division duplex (TDD), and the radio design aims at fulfilling the stringent link requirements.

Table 2-1: sub-THz communication scenarios requirements.

	Mid-range wireless access	**Short-range wireless access**	**Very short-range wireless access**
Example use cases	Digital twins for manufacturing, fixed wireless access, Wireless fronthaul	Digital twins for manufacturing, fully-merged cyber-physical worlds	Fully-merged cyber-physical worlds, holographic communication
Targeted data rate	100 Gbps	10 Gbps	100 Gbps
Typical link range	200 m	10-100 m	10 m
E2E latency	0.1 – 100 ms	0.1 – 100 ms	< 20 ms
Mobility	Stationary (0 m/s)	Mid-speed vehicular (<15 m/s)	Walking speed (<3 m/s)
Radio channel	Outdoor	Indoor/outdoor	Indoor/outdoor
Device classes	AP	AP, mobile device	AP, mobile device
Radio design type	Symmetric	Asymmetric	Asymmetric
Duplex mode	TDD	TDD	TDD
Carrier frequency	140 GHz, 200 GHz, 300 GHz	140 GHz, 200 GHz, 300 GHz	140 GHz, 200 GHz, 300 GHz
Positioning / sensing accuracy	0.1-1 m	0.01 m	<0.01 m
Positioning / sensing latency (depends on mobility)	10 – 100 ms	100 ms	1-100 ms
Delay/distance resolution	0.5 m	0.1 m	0.1 m
Angle resolution	10 degrees	2-10 degrees	2 degrees

Mid-range wireless access: this scenario can be encountered in use cases that require fixed assess with no mobility. The targeted data rate is 100 Gbps at a maximum range of 200 m, which can be realized

by means of high gain antenna arrays. The communicating nodes such as APs have the same capabilities, and therefore, the same radio design can serve for both end nodes.

Short-range wireless access: this scenario requires lower peak data rate, namely 10 Gbps, but mobility should be considered, which requires efforts to cope with the channel variation via beam tracking. When the mobility is controllable in terms of location and trajectory and/or other contextual information (e.g., device position and orientation) is available, such information can be used to predict the beam switching. The maximum range to be covered is 10-100 m in indoor or outdoor environment, with the possibility to exploit spatial multiplexing. Two radio HW needs to be designed differently for the infrastructure AP and the mobile terminal considering the constraints on size, cost, and energy consumption.

Very short-range wireless access: this scenario is defined by 100 Gbps peak data rate and a very short distance - up to 10 m – for providing connectivity to mixed reality and holographic communications with E2E latency less than 20 ms depending on the application. The corresponding channel is indoor and outdoor, with mobility at walking speed. For different constraints on the AP and user equipment, as in the short-range scenario, two different radio HW are needed.

2.2 Flexible radio hardware design

Flexible radio design aims at having different radio architectures for performing different tasks. Several radio options will be designed according to given performance metrics, device constraints, and operating band. In this regard, performance modelling of different radio designs for different scenarios is essential to select a solution that optimizes the cost and energy consumption. This requires a solid methodology and definition of interfaces to the network designer, where the link performance metrics are provided, such as achievable data rate, range, power consumption, latency, and reliability. The network planning and operation can then select the radio options and deployment strategy that fulfil the E2E requirements. Multiple radio options contribute to the network flexibility at the network planning and during the operation. The objective of such design is to fulfil certain link performance metrics with high energy efficiency. The design analysis provides concrete parameters for the generic hardware architecture shown Figure 2-1. This includes the number of antennas, number of signal chains, beamforming network, the required output power per power amplifier (PA), and the noise figure per low-noise amplifier (LNA). This analysis takes also into consideration the hardware impairments models, the impact of waveform, and the channel conditions.

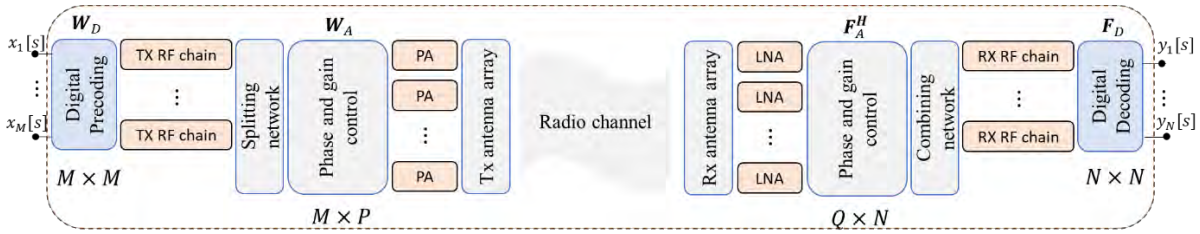


Figure 2-1: Generic radio architecture.

2.2.1 Radio hardware options

In theory, the radio optimization can be performed per specific operating conditions. However, in practice such possibility is not practical as the HW manufacturing complexity needs to be considered. Therefore, the focus is on achieving few radio options with sufficient adaptability optimized for certain ranges. Although a design that covers long range can serve for shorter ranges as well, the cost can be high. Thus, such flexibility will be provided as an option for the vendors and operators to decide how to implement and deploy the radios. In this work, optimal radio designs are investigated per AP and mobile devices. Thereby, there will be in total 5 different radio designs per frequency band, as one radio is not able to operate in all bands. For the performance evaluation, 28 GHz will be considered as a reference, in addition to 140 GHz (D-band), as well as 200 GHz. A sketch of possible exploitation of

the radio options in RAN is shown in Figure 2-2, which highlights the three scenarios and related use cases.

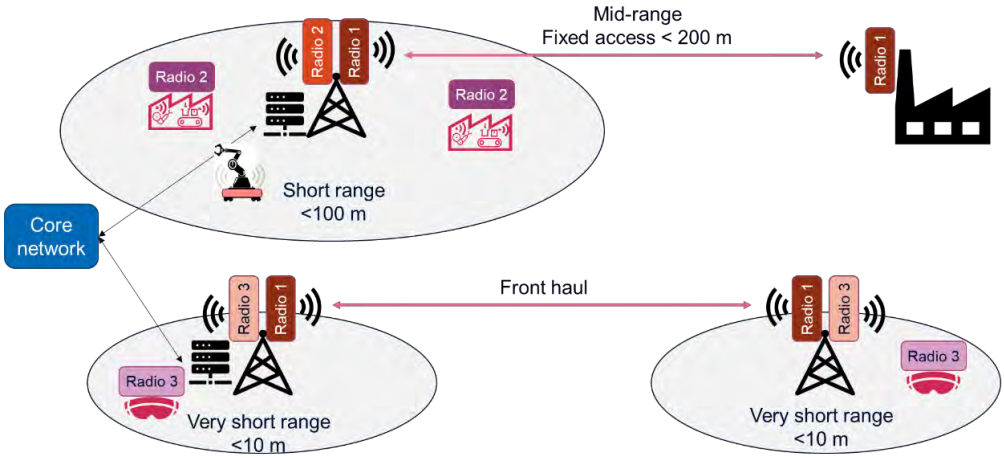


Figure 2-2: Deployment scenarios of different radio hardware options.

2.2.2 Radio hardware design methodology

Starting from the targeted peak data rate and the communication ranges, a first iteration aims at defining initial parameters for the transmitter and receiver to achieve the required SNR budget in a basic line-of-sight (LoS) link estimation. In the next step, the initial design needs to be tested in a particular scenario defined by the environment, and the location of the transmitters and receivers, to evaluate the impact of the radio channel. This analysis aims at finding the number of possible paths for potential exploitation of spatial multiplexing and to determine the degradation of the SNR or the violation of the KPI constraints, such as reduced data rate at the required range, or reduced range for a given data rate. The channel measurements performed in a certain environment can also be used for the evaluation of the digital signal processing (DSP) algorithms, such as channel estimation and equalization. Based on this performance evaluation, the parameters of the radio design need to be tuned. As there can be different realizations, for instance, either by changing the number of antennas or the output power, further refinement is required to select a solution that is practically feasible and minimizes the overall power consumption. The corresponding optimization problem involves many parameters that are constrained by the feasibility and practical limitations. For example, considering the maximum output power of PA, the antenna array size, the resolution of analogue-to-digital converter (ADC), etc. In the context of ADC, the split of the ultrawide band into multiple narrowband channels, which is similar to the analogue implementation of multicarrier system, provides another degree of freedom that allows reusing conventional baseband HW in parallel [HEX21-D22].

To handle the complexity of multiple variables, one approach is based on fixing particular parameters or bounding them to a finite space and fine-tuning others. For instance, one waveform can be selected, and the radio design is tuned to work with this waveform. The radio architecture will be then used for the evaluation of the developed DSP algorithms in practical network settings. At this stage, other important parameters from the use cases need to be considered – in particular latency, which depends on the frame structure and the processing latency. While also link reliability has an impact, a sufficient margin in the SNR budget should be included to reduce the number of retransmissions.

Following the methodology used in [HEX21-D21] and [HEX21-D22], and by elaborating the HW models and considering the channel measurements of the relevant deployment environments, the radio design steps are summarized in the following steps, and further analysis related to the radio HW architecture, and its requirements are described in Section 4.1. A flow chart is shown in Figure 2-3

1- <u>Determine the required SNR condition:</u> from the data rate of the scenario and the available bandwidth, it is required to estimate the required SNR rate using information theory or based on the modulation and coding schemes performance measure such as the maximum symbol

error rate (SER) computed by the Q-function. In this step, the impact of the number of possible spatial streams is provided.

2- Link model: this consists of the transmitter and receiver SNR models, considering the HW impairments that depend on the signal, such as PA and LNA non-linearities, phase noise, quantization noise, and linear distortions such as IQ imbalance and filtering, in addition to the additive noise, which depends on the bandwidth and noise figure. Moreover, the impact of DSP, waveform, and other communication overhead should be considered. The receive signal power depends on the transmit power and the link budget. The link budget is a function of range, carrier frequency, and antenna array and antenna element gains, and it is affected by the path losses of the propagation environment. Regulatory requirements may limit the maximum total transmit power and the spectral mask of the transmitted signal that needs to be considered. In addition, it is necessary to include margins to account for unforeseen or unlikely effects during the system design. This model involves many parameters, and some of them are bounded by the HW limitations, such as the PA transmit power, the antenna array size and number of elements constraints. These constraints need to be considered in limiting the number of possible solutions. Another factor to be considered is to minimize the power consumption and the effective isotropic radiated power (EIRP) radiation, in addition to reducing the cost.

3- Evaluation under radio realistic channel: the parameters obtained need to be evaluated using channel measurements to determine the possible number of independent spatial directions. In theory, this evaluation requires accessing full channel measurements between all pairs of antennas to study the channel rank seen by the antenna array. Each singular value represents the path gain of a spatial direction/beam. However, in a practical sense, obtaining the spatial information can be achieve by assessing several directions, each served by a subarray. High number of independent directions/beams allows the adaptation of transmit power or reducing the bandwidth. On the contrary, in the case of a single non-line-of-sight (NLoS) beam, the link budget may deteriorate, and thus, an update of the design is required. Therefore, an extensive simulation covering different locations determines whether the design requires changes or it can be used in an adaptive way.

4- Evaluation of DSP algorithms: in order to validate the overall air interface design, the selected radio architecture should be used to evaluate the selected waveforms and verify the ability of achieving reliable transmission within the latency constraints of the use cases. In addition, beamforming procedures for initial access and management should be evaluated on top of the selected architectures for the corresponding scenarios.

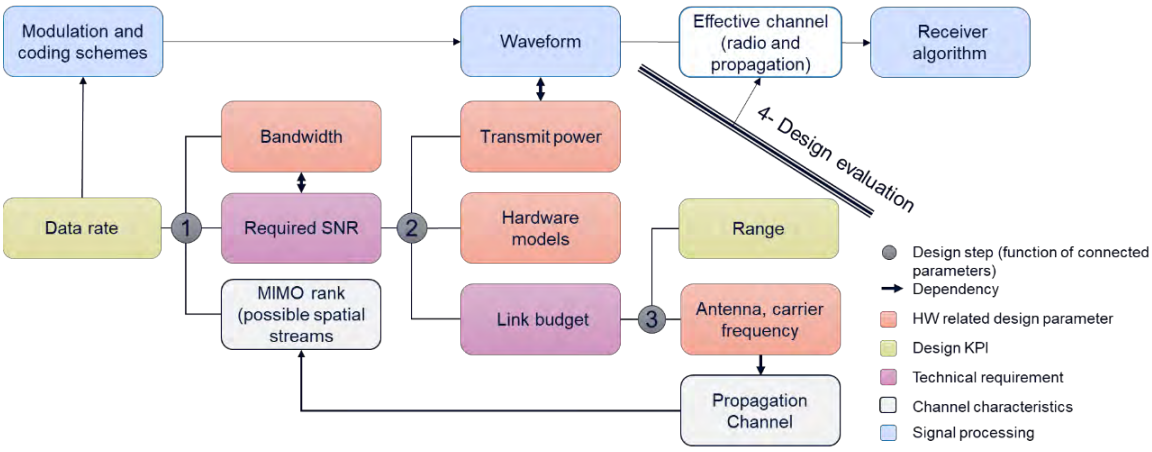


Figure 2-3 Radio design flow chart.

2.3 Performance metrics

For the evaluation of the performance of the selected radio design, several metrics can be used with realistic channel measurements and deployment scenario. The performance metrics will be delivered to the network planning for the study of optimal deployment.

Coverage: this metric provides the coverage area of an AP, where the link data rate at the maximum range is achievable. The maximum range defines the boundary of the coverage. The coverage can be evaluated via the achievable SNR or error vector magnitude (EVM).

Energy performance: is defined as the energy consumption required to fulfil a set of performance requirements. Energy performance is preferred over other commonly used energy related metrics such as *energy consumption* and *energy efficiency*. Minimizing *energy consumption* without considering performance requirements is trivial (just turn the system off). Maximizing *energy efficiency*, e.g. in terms of the required energy to transmit a bit at both the mobile device and access point side. By focusing on *energy performance,* it becomes clear that performance requirements might be also managed to reduce energy consumption. Requirements have a cost in terms of energy usage, and requirements supporting advanced services only need to be fulfilled when and where these services are used.

Link reliability and latency: this metric concerns the actual evaluation of the communications system in terms of block error rate (BLER) that needs to be fulfilled within the latency budget, considering also the communication overhead, under the assumption of an established link. The reliability can be defined as the number of the correctly conveyed frames within the air interface latency budget to the total number of communicated frames. Accordingly, processing latency, frame structure, and buffering delays need to be considered.

Beam failure rate: this measure determines the connection breaks per timer interval. The connections breaks are especially due to mobility when the beam tracking fails to keep the connection. Note that this can also happen during the initial beam acquisition.

Positioning performance metrics: a comprehensive list of localization and sensing related performance metrics can be found in [HEX21-D31, Table 2.1]. Among there metrics are

- Positioning accuracy (2D or 3D): root mean square (RMS) value of the location error in metres. Sometimes the error is defined in vertical (perpendicular to the earth surface) and horizontal (parallel to earth surface) instead of a 3D accuracy.
- Orientation estimation accuracy (3D): RMS value of the orientation error in a global frame of reference. Typically measured in rad. Alternatively, accuracy can be measured as the error value corresponding to a certain percentile.
- Positioning latency: The duration between an application or a network function requesting location and obtaining the results. Includes the physical layer latency.

These metrics depend on the accuracy of the estimated parameter used in computing the position and orientation including, time-of-arrival (ToA), angle-of-arrival (AoA), and angle-of-departure (AoD), which are influenced by the radio design and HW constraints.

3 Channel model

This section describes channel models and their input parameters that were populated in the project. In order to support various types of channel models ranging from fully physical model, e.g., ray-optics simulations, to fully stochastic channel model based on the geometry, two basis channel models are provided: 1) material parameters for 2-260 GHz and 2) stored channel model at 140 GHz that uses processed measured channel responses as a seed to reproduce channels based on a Monte Carlo simulation.

3.1 Material permittivity and conductivity estimation from 2 to 260 GHz

Physical models such as ray tracing or ray launching models have been using a geographical database to compute all possible paths (direct, reflected, diffracted, diffused paths) between two or more points such as transmitter (Tx), receiver (Rx), reconfigurable intelligent surfaces (RIS), relay, etc. These models are consequently frequently used for the performance evaluation of new wireless radio interfaces. The path amplitude is given by physical laws according to the interaction type between the electromagnetic (EM) wave and the environment. For instance, the direct path amplitude is given by the well-known free space loss equation, the diffracted path amplitude is given either by the knife-edge diffraction theory or by the uniform/geometry theory of diffraction (UTD/GTD), and the reflected or transmitted path amplitude is given by the Fresnel equations. Compared to losses due to other physical phenomena, reflection and transmission losses depend on two material properties, the permittivity and conductivity, that may be frequency dependent. Hence, a physical model shall include a material table indicating the material permittivity and conductivity and, if necessary, how these parameters change with frequency. Many simulations refer to the ITU model [ITU P.2040-2] that is valid up to 100 GHz. The ITU model may not be valid for 6G simulations where sub-THz frequencies up to 300 GHz are considered. The goal of this section is to discuss the ITU model validity for frequencies above 100 GHz. The analysis is based on continuous material reflection and transmission loss measurements from 2 to 260 GHz.

3.1.1 Theoretical model and ITU model

Let μ_0 and ε_0 be the vacuum permeability and permittivity, respectively, and ε_r and σ the dielectric relative permittivity and conductivity, respectively. In usual human environments, EM waves travel from a transmitter to a receiver located in the air and interact with building materials that may be considered as flat-surface homogeneous isotropic low-loss dielectrics with $(\sigma/\omega\varepsilon_0\varepsilon_r)^2 \ll 1$. Building materials are usually represented by a slab with thickness *d* as indicated in Figure 3-1. The slab may be a glass window, concrete wall, plasterboard partition wall, etc. Figure 3-1. illustrates the different paths created by the interaction between the EM wave and the slab. As the propagating wave reaches the first material interface, a part of the signal is reflected, and the rest is transmitted through the material. The transmitted part from the first material interface penetrates through the material, and it is attenuated due to the penetration loss. Arriving to the second interface, a part of the signal will be transmitted, and the rest will be reflected. Likewise, multiple reflections are created inside the slab. In the following, for the sake of simplicity, only transmitted paths 1 and 2 (named 1T and 2T) and reflected paths 1 and 2 (named 1R and 2R) will be considered, as illustrated in Figure 3-1. The frequency-dependent complex reflection $H_r^{Mod}(f)$ and transmission coefficients $H_t^{Mod}(f)$ of the material slab are expressed as $H_r^{Mod}(f) = r - \sqrt{\varepsilon_r}t^2p^2t$ and $H_t^{Mod}(f) = \sqrt{\varepsilon_r}t^2p - \sqrt{\varepsilon_r}t^2p^3r^2$. Here, r and t are the reflection and transmission coefficients at the air-material interface . At a normal incidence, $r = \frac{1-\sqrt{\varepsilon_r}}{1+\sqrt{\varepsilon_r}}$ and $t = \frac{2}{1+\sqrt{\varepsilon_r}}$. Here, p is the penetration loss inside the material, and α and β are the attenuation and phase constants and can be approximated by $\frac{\sigma}{2}\sqrt{\frac{\mu_0}{\varepsilon_0\varepsilon_r}}$ and $\omega\sqrt{\mu_0\varepsilon_0\varepsilon_r}$, respectively.

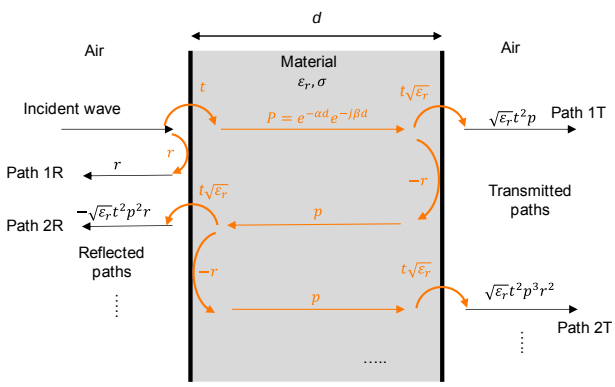


Figure 3-1: Multipaths inside a material slab.

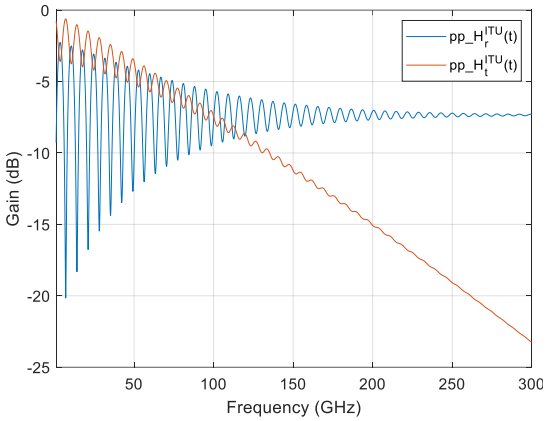


Figure 3-2: $pp\_H_r^{ITU}(f)$ and $pp\_H_t^{ITU}(f)$ for glass.

The ITU recommendation [ITU P.2040-2] proposes a constant permittivity value regardless the frequency and expresses the frequency-dependent conductivity $\boldsymbol{\sigma}$ as $\boldsymbol{\sigma} = \boldsymbol{cf^d}$. $\boldsymbol{\varepsilon_r}$, c and d are given for some materials such as glass. $H_r^{ITU}(f)$ and $H_t^{ITU}(f)$ are $H_r^{Mod}(f)$ and $H_t^{Mod}(f)$ applying the ITU model. $H_r^{ITU}(f)$ and $H_t^{ITU}(f)$ are illustrated in Figure **3**-2 for the glass ($d = 8$ mm, $\varepsilon_r = 6.25$, $\sigma = 0.0043 f^{1.19}$).

When analysing propagation channels in indoor or outdoor environments, the large-scale effects are defined as referring to the environment or Tx/Rx distance effects, while the small-scale effects refer to the multipath complex combination. It can be observed in Figure 3-2 similar effects with a large-scale frequency variation that depends on the permittivity and conductivity. The higher the permittivity, the lower the reflection losses. The higher the conductivity, the higher the penetration losses. Small-scale effects are created by the interferences between the multipath created inside the slab. Due to the penetration loss increase with the frequency, small-scale effects tend to become less relevant at high frequencies.

As a first approach, the discussion about the ITU model validity above 100 GHz may be divided in two sub-discussions. The first sub-discussion is about the validity of the ITU theoretical framework. For instance, does the permittivity remain constant and is the conductivity modelled by cf^d ? Are the conductivity and permittivity still relevant parameters to compute reflection and transmission losses? The second sub-discussion is about the numerical values of ε_r, c and d. Shall they be updated to correctly represent reflection and transmission losses? An ITU-like model was defined to help the analysis related to the first sub-discussion. This model is called ITUF and keeps the same theoretical approach of the ITU model but parameters ε_r, c and d are fitted on measurement. The notations $H_r^{ITUF}(f)$ and $H_t^{ITUF}(f)$ correspond to the transfer functions $H_r^{Mod}(f)$ and $H_t^{Mod}(f)$ when applying the ITUF model. A pp prefix added to H functions indicates the power profile of H functions with $pp\_H(f) = 10 * \log_{10}(|H(f)|^2)$.

3.1.2 Measurement description and data processing

The measurement equipment is based on a vector network analyser (VNA) and frequency extenders. It can measure S parameters continuously from 2 to 260 GHz. The measurement is divided into contiguous sub-bands defined by the frequency limits of the VNA, extenders, or antennas (2-30 GHz, 30-50 GHz, 50-75 GHz, 75-110 GHz, 110-170 GHz, 170-260 GHz). Antennas are dual-ridge ultrawideband (UWB) antennas for frequencies below 50 GHz and standard horn pyramidal antennas for frequencies above 50 GHz. Figure 3-3 illustrates the mechanical part that allows a free space measurement for both transmission and reflection coefficients at normal incidence. The material under test was placed at normal incidence between the two antennas in the material holder middle. Figure 3-3 gives some slab examples.

Common materials in an indoor environment were measured. They are grouped in category depending on the material construction properties or the material function:

- Glass: coated glass, laminated glass, patterned glass
- Plasterboard: BA13, BA18, water and humidity resistant BA18 (W&H BA 18)
- Wood and its derivatives: plywood, medium density fibreboard (MDF), melamine chipboard.
- Concrete-like materials: mortar, cellular concrete
- Flooring tiles: vinyl, ceramic, carpet
- Thermal insulation materials: polystyrene, glass wool ceiling tile
- Miscellaneous: plexiglass, gypsum fibre board (GFB), and plastic fibre board (PFB).

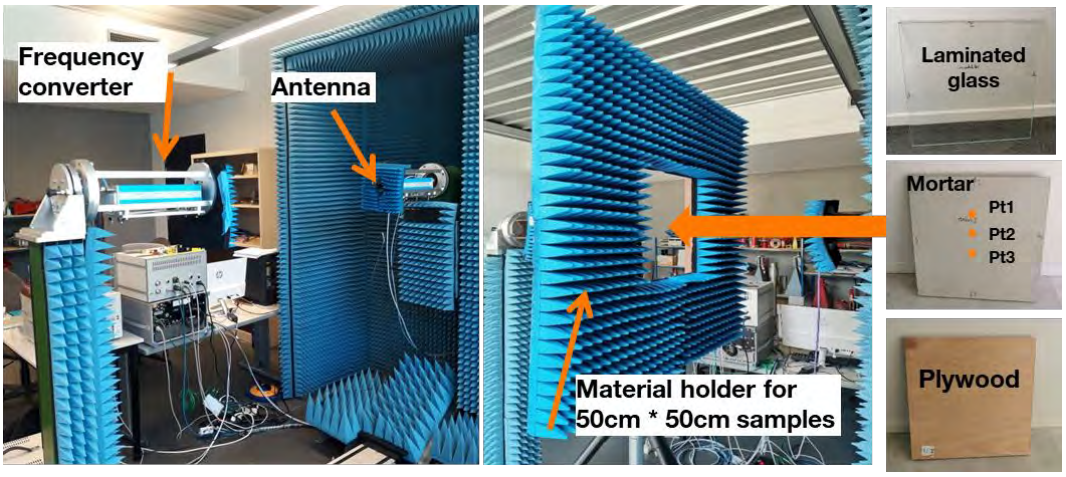


Figure 3-3: Measurement system.

For each frequency sub-band, the material measurement is preceded by the reference measurement in free space or with a metallic plate instead of a material. The terms $S_{11}^{FS}(f)$ and $S_{21}^{FS}(f)$ refer to the S parameters measured by the VNA in free space, and $S_{11}^{Met}(f)$ the S11 parameter measured when inserting a metallic plate into the material holder. When reference measurements are completed, slabs are put in the material holder and material measurements are proceeded. The measurement is performed at 3 points separated by 10 cm. The raw material measurements are denoted by $S_{11}^{Meas}(f)$ and $S_{21}^{Meas}(f)$ for reflection and transmission measurements, respectively. Moreover, $H_r^{MUT}(f)$ and $H_t^{MUT}(f)$ represent the intrinsic material transmission and reflection frequency response. They are computed according to (3-1),

$$H_r^{MUT}(f) = \frac{TG\left[S_{11}^{Meas}(f) - S_{11}^{FS}(f)\right]}{TG[S_{11}^{Met}(f) - S_{11}^{FS}(f)]} \qquad H_t^{MUT}(f) = \frac{TG[S_{21}^{Meas}(f)]}{TG[S_{21}^{FS}(f)]}, \tag{3-1}$$

where $TG[.]$ is the time gating operator. It is a temporal filter applied on the inverse fast Fourier transform (IFFT) of S parameters that rejects all undesirable components such as wall reflections or material edge diffractions. Here, $h_r^{MUT}(t)$ and $h_t^{MUT}(t)$ denote the IFFT of $H_r^{MUT}(f)$ and $H_t^{MUT}(f)$, respectively. In the following sections, the data analysis and model comparison will be mainly based on power profile. A $pp\_$ prefix added to a complex vector indicates the vector power profile. For instance, $pp\_h_r^{MUT}(t)$ is the power profile of $h_r^{MUT}(t)$ and is equal to $10 * \log_{10}(|h_r^{MUT}(t)|^2)$.

3.1.3 Conductivity/permittivity estimation

As the measurement presented in this work are complex over very large bandwidths, it is possible to resolve, in the time domain, the multipath described in the theoretical part. For instance, Figure 3-4 shows $pp\_h_{11}^{MUT}(t)$ for the plexiglass at different frequency ranges. Path 1R and 2R defined in Figure 3-1 can be clearly identified. X-scale is converted from the time domain to the distance domain by assuming the light speed in vacuum.

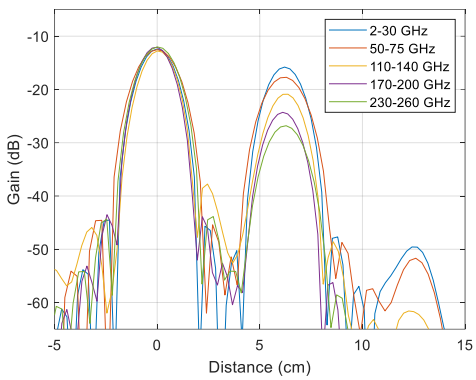


Figure 3-4: $pp\_h_r^{MUT}(t)$ for plexiglass.

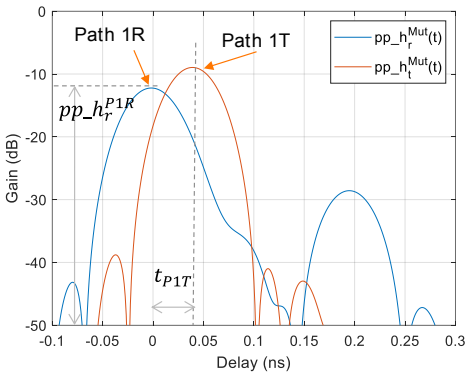


Figure 3-5: Permittivity estimation method.

Consequently, separate estimation methods are applied for the permittivity and conductivity. First, the permittivity is estimated by two different methods based either on the delay of transmitted paths or on the amplitude of reflected paths. Only path 1R amplitude and path 1T delay are considered, as illustrated in Figure 3-5, because second order paths are often below the noise level at the highest frequencies. The initial measurement sub-bands were slightly rearranged in order to define nine new sub-bands of about 30 GHz bandwidth (2-30 GHz, 30-50 GHz, 50-75 GHz, 75-110 GHz, 110-140 GHz, 140-170 GHz, 170-200 GHz, 200-230 GHz, 230-260 GHz). An analysed bandwidth of roughly 30 GHz is a good trade-off to keep a good resolution for the path detection, to consider a constant value for the permittivity and to assess the large-scale frequency dependency of permittivity.

Method 1 is based on the delay. Let t_{P1T} be the delay of path 1T. The wave speed inside the material is equal to $\frac{c_v}{\sqrt{\varepsilon_r}}$, c_v being the wave speed in vacuum. Then, the relative permittivity ε_r is calculated as expressed in (3-2).

$$\varepsilon_r = \left(\frac{t_{1T} * c_v}{d} + 1\right)^2. \quad \textbf{(3-2)}$$

Method 2 is based on the amplitude. Let's define h_r^{1R} the amplitude of path 1R. $h_r^{1R} = \frac{\sqrt{\varepsilon_r} - 1}{\sqrt{\varepsilon_r} + 1}$, then, ε_r can be calculated as in (3-3).

$$\varepsilon_r = \left(\frac{1 + h_r^{1R}}{1 - h_r^{1R}}\right)^2. \quad \textbf{(3-3)}$$

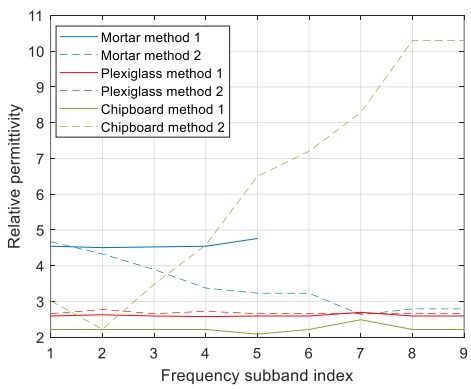


Figure 3-6: Estimation method comparison.

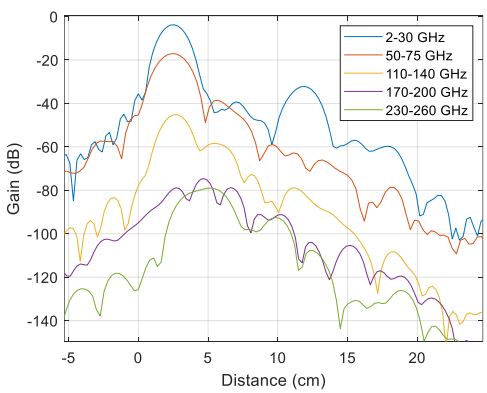


Figure 3-7: $pp\_h_t^{MUT}(t)$ for mortar.

Figure 3-6 shows the permittivity estimated for the plexiglass, mortar, and chipboard. The permittivity estimation is sometimes unreliable for rough surface materials or inhomogeneous materials, or multi-

layer materials because $pp\_h_r^{MUT}(t)$ or $pp\_h_t^{MUT}(t)$ do not show sufficient distinction of path 1R or path 1T. For instance, as illustrated in Figure 3-6, the mortar permittivity was not estimated by method 1 above 140 GHz. Permittivity values estimated by method 1 and method 2 are very similar for most of the materials but diverge for some materials when the frequency increase. It can be observed in Figure 3-6 that the permittivity estimated by method 2 for the mortar decreases artificially due to the scattering from rough surfaces when the wavelength is similar to the surface small irregularities. This point is discussed more in details in the next paragraph. It can also be seen that the chipboard permittivity estimated by method 2 strongly increases with frequency, whereas the permittivity estimated by method 1 remains constant. The permittivity increases by method 2 due to the melamine at the chipboard surface or due to the chipboard composite structure. The small wood particles may become more efficient reflectors when the wavelength decreases.

Even if the permittivity is mainly used to compute reflection losses in propagation channel simulation, the final permittivity values are estimated by method 1, i.e., based on the delay of transmission measurement, because this method seems to be more representative of the material electrical properties and less dependent of the material surface state.

Figure 3-8 presents the permittivity for various materials. It confirms that the permittivity is independent of the frequency even at frequency above 100 GHz. The permittivity is averaged over the different frequency bands, and the mean value is reported in Table 3-1

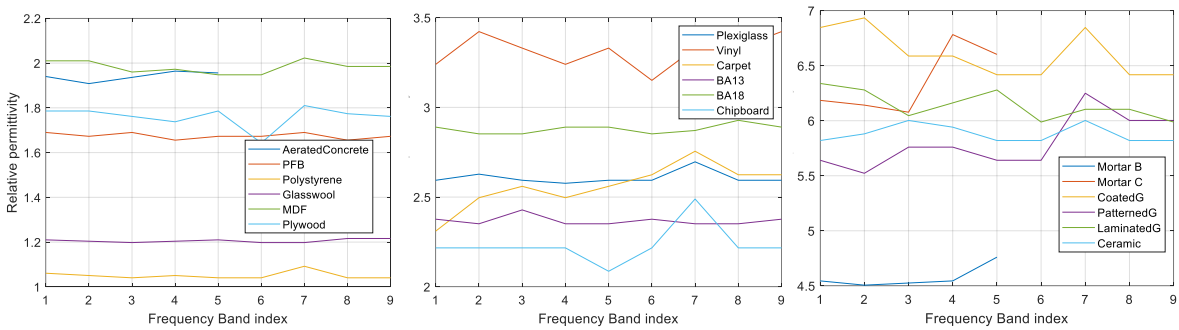


Figure 3-8: Material permittivity.

The conductivity estimation was not performed by detecting path 1T amplitude for each frequency sub-band as the conductivity decreases with frequency and as the transmission losses are highly impacted by the material inhomogeneity. This point is discussed more in details in the next paragraph. Consequently, a simple and robust approach is to estimate parameters c and d by an iterative approach in order to minimize the error between $pp\_H_t^{MUT}(f)$ and $pp\_H_t^{ITUF}(f)$. The conductivity is estimated assuming the constant permittivity as indicated hereabove.

3.1.4 Comparison between measurement and the ITUF model

Figure 3-9 shows $pp\_H_r^{MUT}(f)$ and $pp\_H_t^{MUT}(f)$ for three typical materials: plexiglass, mortar with moderate surface roughness, and melamine chipboard. The experimental results are compared with the ITUF model and when possible, with the ITU model. For clarity reasons, the ITU model is computed by considering only path 1T in transmission and path 1R in reflection.

The ITUF model for the plexiglass agrees very well with $pp\_H_r^{MUT}(f)$ and $pp\_H_t^{MUT}(f)$, even at frequencies above 100 GHz. No difference is observed between the three measurement points. Similar results are found for the different glasses, and it can be concluded that the ITUF model can be used without any restriction when the material is homogeneous with a flat surface. For all other materials, the ITUF model remains valid up to a frequency that depends on the material structure.

Regarding reflection losses, the mortar with moderate roughness is a good example to illustrate the scattering impact on the ITUF performance. It can be clearly observed in Figure 3-9 that the regular periodic small-scale fading disappears at 60 GHz and is replaced by an irregular fading with no correlation between the three points. This effect is due to the scattering on the surface. The fading is no

more due to the complex recombination between path 1R and path 2R but is due to the complex recombination of scattered components. The scattering increases the reflection loss average value and makes the ITUF model inaccurate or not valid as the difference model/measurement may be higher than 10 dBs. It can be seen that, for some materials (MDF, PFB, glasswool, etc), $pp\_H_r^{MUT}(f)$ increases with frequency as illustrated by the melamine chipboard example in Figure 3-9. The $pp\_H_r^{MUT}(f)$ increase compared to $pp\_H_r^{ITUF}(f)$ is generally limited to a few dBs.

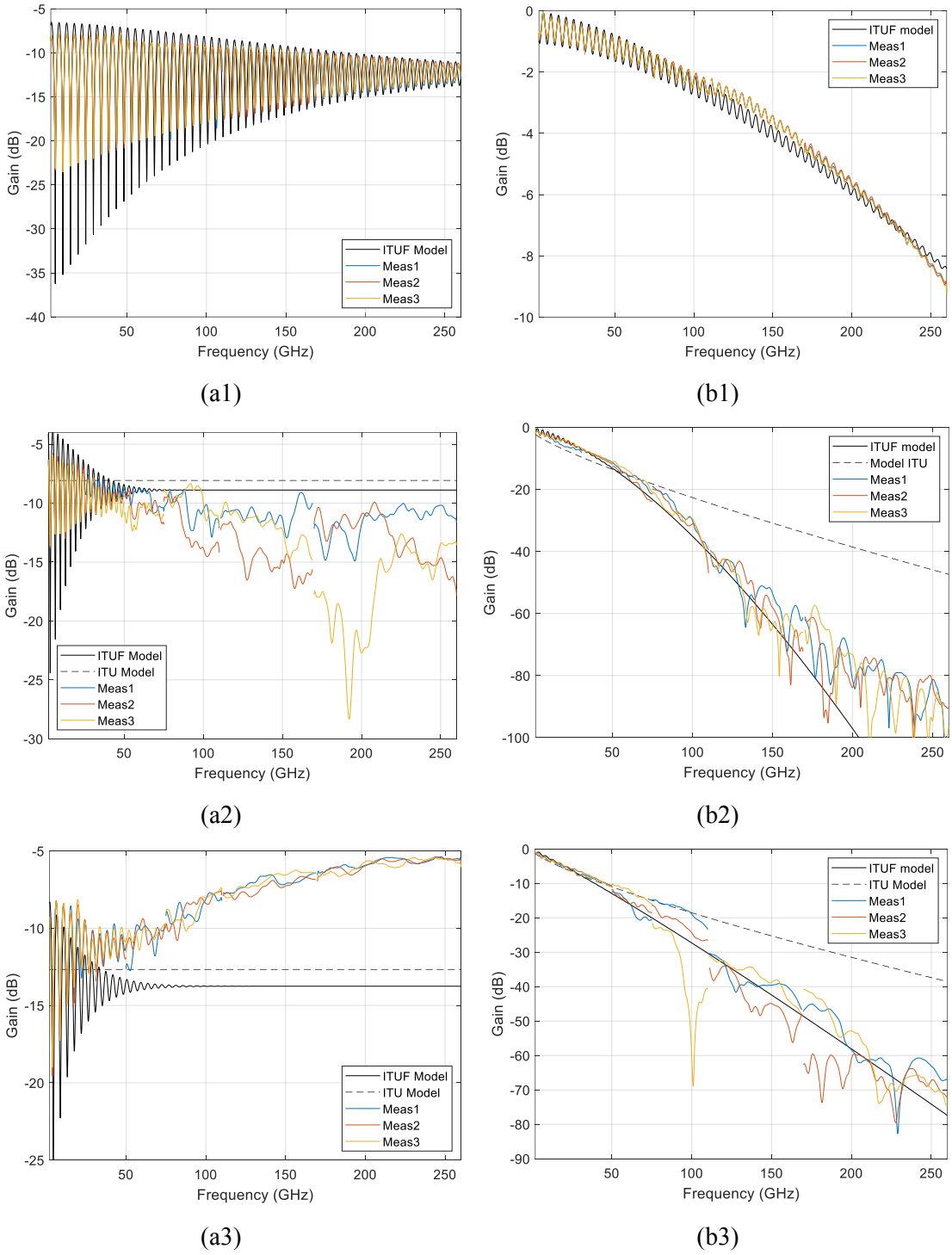


Figure 3-9: $pp\_H_r^{MUT}(f)$ and $pp\_H_t^{MUT}(f)$ compared with ITU and ITUF models. Plexiglass (a1,b1), Mortar (a2,b2), Melamine chipboard (a3,b3).

The analysis of transmission results leads to similar conclusions. The frequency-dependent conductivity can be well modelled by the ITUF model "on average". The absolute difference between $pp\_H_t^{ITUF}(f)$ and $pp\_H_r^{MUT}(f)$ is generally about a few dB for frequencies less than 60 GHz but may exceed 10 dB for frequencies above 100 GHz. For instance, the ITUF model agrees with mortar or chipboard results up to 60 GHz, but some variations appear above 60 GHz. Above 100 GHz, the transmission loss differences between the ITUF model and measurements can exceed 10 dB, and the three points results are uncorrelated. Figure 3-7 shows $pp\_h_t^{Mut}(t)$ for the mortar at different sub-bands. Path 1T is highly attenuated, and some delayed components appear at frequencies above 100 GHz. It suggests that the volumic diffuse scattered component inside the material due to the material inhomogeneity becomes dominant over the specular transmission component.

3.1.5 ITU model validation above 100 GHz

The estimated permittivity and conductivity, which are based on measurements, are reported in Table 3-1 and compared with the ITU model values when available. There is a good agreement between the experimental values and those provided by ITU. Consequently, the permittivity ITU model can be extended for frequency up to 300 GHz. But many building materials are composite or multi-layered materials. When the wavelength decreases, these materials cannot still be considered as homogeneous with flat surface, and the reflection loss may be incorrectly modelled when using the permittivity-dependent Fresnel equations. Applying scattering models is one option but is complex to implement in ray tracing and is limited to rough surfaces. Above 100 GHz, a simple model evolution could be the definition of correction factors to consider the reflection loss decreases or increase compared to the Fresnel reflection coefficients. This approach would require new measurements to investigate more in details the reflection loss of composite materials or covered material (melamine, wallpaper, painting, varnish, etc).

Table 3-1: Estimated and ITU permittivity and conductivity.

		Estimated value ε_r	ITU model ε_r	Estimated value		ITU model	
				c	d	c	d
Glass	Laminated glass	6.2	6.27	0.005	1.2	0.0043	1.1925
Plasterboard	BA13	2.5	2.94	0.0004	1.35	0.0116	0.7076
	BA18	2.9		0.0004	1.4		
Wood and derivatives	MDF	2.0		0.005	1.0		
	Plywood	1.8	1.99	0.006	1.0	0.0047	1.0718
	Chipboard	2.25	2.58	0.007	1.1	0.0217	0.78
Concrete-like materials	Aerated concrete	1.9		0.002	1.3		
	Mortar 1	4.6	5.31	0.002	1.5	0.0326	0.8095
	Mortar 2	6.4		0.004	1.4		
Flooring	Floor vinyl tile	3.3		0.0002	1.6		
	Floor ceramic tile	5.9		0.005	1.2		
	Office carpet	2.5		0.005	1.05		
Insulation plates	Polystyrene	1.05		0.000008	1.1		
	Glass wool tile 1	1.25	1.5	0.0002	1.35	0.0005	1.1634
	Glass wool tile 2	1.2	1.5	0.002	1.2		
Miscellaneous	Plexiglass	2.65		0.0001	1.5		
	GFB	3.45		0.002	1.25		
	PFB	1.7		0.0001	1.55		

Estimated parameters *c* and *d* agree with those given by the ITU for the glass and plywood but are sometimes very different. For instance, no values of *d* less than 1 were observed. It implies that the ITU model underestimates the penetration loss at high frequencies for materials having a *d* value less than 1. Figure 3-9 illustrates the model error for the chipboard and mortar. The recommendation is to consider the dielectric parameters estimated by our measurement at least for the mortar and chipboard. Regarding the transmission loss variation around the model due to the inhomogeneity at high

frequencies, the ITU model could be improved by adding to $H_t^{ITU}(f)$ a statistical variable but as indicated hereabove for the reflection loss, this approach requires additional measurements

Finally, it can be concluded that the ITU model needs to be improved for frequencies above 100 GHz to represent the measured reality of reflection and transmission well, while it 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 office environment, shopping mall, airport, etc as most of the material are quite smooth. Transmission loss errors due to the material inhomogeneity may not be a concern as they concern high transmission losses at high frequencies. Simulating a transmission error of 20 dB instead of 30 dB may not significantly impact system simulation if a blockage is defined by a transmission loss higher than 20 dB. Future work will be focused on multi-frequency measurement in real environments to better evaluate the propagation channel differences between cmW and sub-THz frequencies.

3.2 Stored channel model at 140 GHz

3.2.1 Principle

The use of measured channel responses for physical layer design and evaluation on a computer has been a well-recognized approach, e.g., [MST+02], as it allows repeatable tests and comparison of physical layer 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 several drawbacks of using measured channels for physical layer studies, i.e., 1) measurements are always subject to uncertainties; 2) it is not straightforward to apply the measured channel responses to simulations that assume different hardware characteristics than the measurement, e.g., signal dynamic range, antennas/arrays, system bandwidth and moving speed of a mobile; 3) a sufficient amount of measurements for, e.g., Monte-Carlo simulations and evaluation of packet error rates, may not be available due to limited amount of data that channel sounding can obtain. The measurements are moreover limited to certain environments where channel sounders can be installed; for example, it is not easy to perform Terahertz channel sounding for ground-to-unmanned aircraft scenario. There exists a need to overcome the aforementioned drawbacks by, e.g., i) properly performing calibration of the channel sounder and filtering noise from measured channels, ii) under-sampling or interpolating the measured channels in space, bandwidth and time and iii) by making a general 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 by Molisch *et al.* [MST+02] where band-/aperture-/time-sampling-limited measurements of channels are approximated by the band-/aperture-/time-sampling unlimited form. 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.

Figure 3-10 illustrates the band-/aperture-limited and band-/aperture-unlimited forms of 140 GHz indoor channels. It is possible to reproduce infinite amount of small-scale fading realizations of channels from the propagation paths of channels by applying the uniform randomly distributed phases to each path before summing them up at the antenna. The uniformly distributed phase variation to each path is a well-justified means to emulate the statistics of the random local movement of a mobile terminal. The exact paper [MST+02] analyses the Ergodic channel capacity of MIMO channels from limited MIMO channel sounding, whereas e.g., [SNL+21] uses the same approach to identify the most suitable fading distribution at 140 GHz.

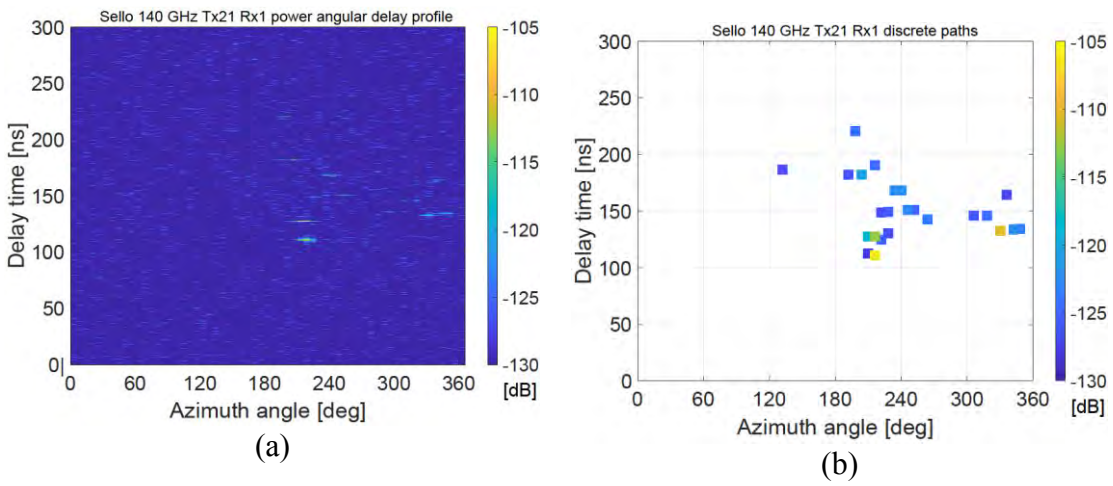


Figure 3-10: (a) Band- and aperture-limited channel response from a measurement and (b) its band- and aperture-unlimited model as propagation paths. Data are from a shopping mall measurement at 140 GHz [NJK+18].

3.2.2 Stored channels from measurements

Channel sounding campaigns were performed in an entrance hall, and outdoor scenarios such as suburban, residential, and city centre. The antenna heights, number of LOS and NLOS links, and link distance range for each environment are summarized in Table 3-2. The measurement covered from 140 GHz to 144 GHz, resulting in 4 GHz bandwidth. The channel sounder is equipped with an omnidirectional biconical antenna on the Tx side and a horn antenna, mounted on a rotator with 10° azimuth half power beam width (HPBW) on the Rx side. The rotator scans the azimuth directions of the horn while the main lobe of the antenna is fixed at the horizontal plane. With this kind of directional measurement, single-directional multipath components denoted as

$$P_q(\varphi^{\mathrm{rx}}, \tau) = \sum_{l=1}^{L_q} P_{q,l}\,\delta\left(\varphi^{\mathrm{rx}} - \varphi_{q,l}^{\mathrm{rx}}\right)\delta\left(\tau - \tau_{q,l}\right), \tag{3-4}$$

were obtained for each radio link; q is the link index, L_q is the number of paths, $\delta(\cdot)$ is the delta function, $P_{q,l}$, $\varphi_{q,l}^{\mathrm{rx}}$, and $\tau_{q,l}$ are the power (squared magnitude of path gain), the azimuth direction of arrival, and the propagation delay of the lth path, respectively. The missing angular knowledge of multipath components, e.g., azimuth angles on the Tx side and elevation angles at the Rx, was supplemented by fusing the available multipath parameter estimates, i.e.,(3-4), with the available detailed geometric database of the measurement environment [DKH21]. The resulting channel data, which are published in [DHK23], consist of double-directional multipath components denoted as

$$P_q(\Omega^{\mathrm{rx}}, \Omega^{\mathrm{tx}}, \tau) = \sum_{l=1}^{L_q} P_{q,l}\,\delta\left(\Omega^{\mathrm{rx}} - \Omega_{q,l}^{\mathrm{rx}}\right)\delta\left(\Omega^{\mathrm{tx}} - \Omega_{q,l}^{\mathrm{tx}}\right)\delta\left(\tau - \tau_{q,l}\right), \tag{3-5}$$

where $\Omega_{q,l}^{\mathrm{rx}}$ and $\Omega_{q,l}^{\mathrm{tx}}$ are the direction of arrival and departure, including azimuth and elevation angles. Current data are measured with vertically polarized antennas; hence the definitions are restricted here to a single polarized case only, neglecting the polarization characteristics.

Table 3-2: 140 GHz spatio-temporal channel sounding.

	Entrance Hall	Suburban	Residential	City Centre
Number of LOS Links	5	5	5	5
Number of NLOS Links	12	32	13	19
Link Distance Range (m)	3-66	2-172	20-175	10-178

3.2.3 Channel generation

From the stored multi-path data described in the previous section, it is possible to generate radio channel responses at 140 GHz with different types of transmit and receive antennas and antenna arrays. Receiver and transmitter antennas can be specified by complex radiation patterns $\mathbf{g}_{\mathrm{rx}}(\Omega^{\mathrm{rx}}) \in \mathbb{C}^{M\times 1}$ and $\mathbf{g}_{\mathrm{tx}}(\Omega_l^{\mathrm{tx}}) \in \mathbb{C}^{N\times 1}$, respectively, where M is the number of Rx antennas and N is the number of transmitter antennas. Now the channel frequency response matrix can be determined as

$$\mathbf{H}_q(f) = \sum_{l=1}^{L_q} \mathbf{g}_{\mathrm{rx}}\left(\Omega_{q,l}^{\mathrm{rx}}\right) \sqrt{P_{q,l}}\, e^{-j2\pi f \tau_{q,l}}\, \mathbf{g}_{\mathrm{tx}}\left(\Omega_{q,l}^{\mathrm{tx}}\right)^T \in \mathbb{C}^{M\times N}. \quad \textbf{(3-6)}$$

Random snapshots of frequency response matrices can be generated by introducing random initial phase $\varphi_{q,l}$ 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

$$\mathbf{H}_q(t,f) = \sum_{l=1}^{L_q} \mathbf{g}_{\mathrm{rx}}\left(\Omega_{q,l}^{\mathrm{rx}}\right) \sqrt{P_{q,l}}\, e^{j(\varphi_{q,l}+2\pi\nu_{q,l}t)}\, e^{-j2\pi f \tau_{q,l}}\, \mathbf{g}_{\mathrm{tx}}\left(\Omega_{q,l}^{\mathrm{tx}}\right)^T \in \mathbb{C}^{M\times N}. \quad \textbf{(3-7)}$$

Examples of band-/aperture-unlimited double directional power angular delay profiles (PADPs) are shown in Figure 3-11 and Figure 3-13. Three colours indicate three propagation paths of three Tx/Rx locations measured at 140 GHz in an indoor environment. Figure 3-11 illustrates power delay profiles (PDP) of measured propagation channel on three different Tx/Rx locations. One can observe that within examples, the maximum excess delay varies between 60 and 200 ns. Moreover, e.g., the link number 3 (blue) has 12 multipath within 20 dB dynamic range. Figure 3-13 depicts power-angular information of the same paths. The top left figure indicates azimuth arrival/departure angles, top right elevation arrival/departure angles, bottom left arrival azimuth/elevation angles, and bottom right departure arrival/departure angles. In the illustrated examples, the elevation and azimuth angles are within approximately 40 deg and 220 deg range, respectively.

Another source of time variability is the blockage of a propagation path by a moving obstacle, e.g., a human body. Such events are probable, especially in short link distance indoor environments, when antennas are not elevated high and persons are moving in the environment. Short wavelengths at sub-THz radio frequencies make Fresnel ellipsoids of each multipath component small and a human body can severely block the signal up to 40 dB attenuation. The human blockage effect was measured at 140 GHz frequency at 3.5 m link distance with different persons and orientations of persons in [ZBK+23].

Blockage events are modelled as supplemental time variability in (3-7) by introducing attenuation term $\alpha(t)$ such that the resulting transfer function becomes

$$\mathbf{H}_q(t,f) = \sum_{l=1}^{L_q} \mathbf{g}_{\mathrm{rx}}\left(\Omega_{q,l}^{\mathrm{rx}}\right) \sqrt{\alpha_{q,l}(t) P_{q,l}}\, e^{j(\varphi_{q,l}+2\pi\nu_{q,l}t)}\, e^{-j2\pi f \tau_{q,l}}\, \mathbf{g}_{\mathrm{tx}}\left(\Omega_{q,l}^{\mathrm{tx}}\right)^T \in \mathbb{C}^{M\times N}. \quad \textbf{(3-8)}$$

The attenuation term is taken directly from the measurement reported in [ZBK+23]. Figure 3-12 shows two example attenuation patterns, one with lateral and one with frontal crossing of a human body across a multipath component. In the channel model one can choose which path or paths are blocked, the initial time instance of each blockage event, and the type of blockage, i.e., lateral or frontal crossing of a human body. These choices can be made also random. If path l experiences blockage, the term $\alpha_{q,l}(t)$ gets such values as in Figure 3-12 in linear units, starting with the selected time instant. Term $\alpha_{q,l}(t) = 1\ \forall t$ are used for those paths without blockage event. A few remarks can be made from Figure 3-12 about the human shadowing at short link distances. The maximum attenuation reaches approximately 40 dB. The frontal blocking has shorter duration and slightly less severe attenuation, as expected. There are diffraction patterns just before after the actual blockage. Potentially these power variations could be used to predict an upcoming blockage event.

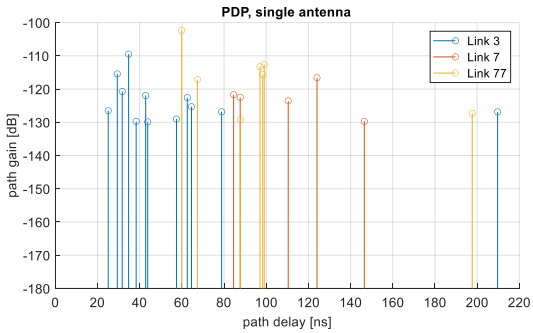


Figure 3-11: Power delay profiles of three example links of the stored channel model.

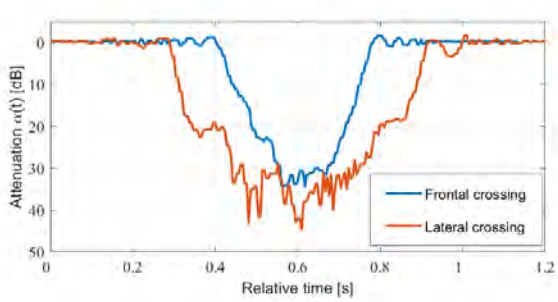


Figure 3-12: Time variant attenuation pattern of frontal and lateral human blockage events.

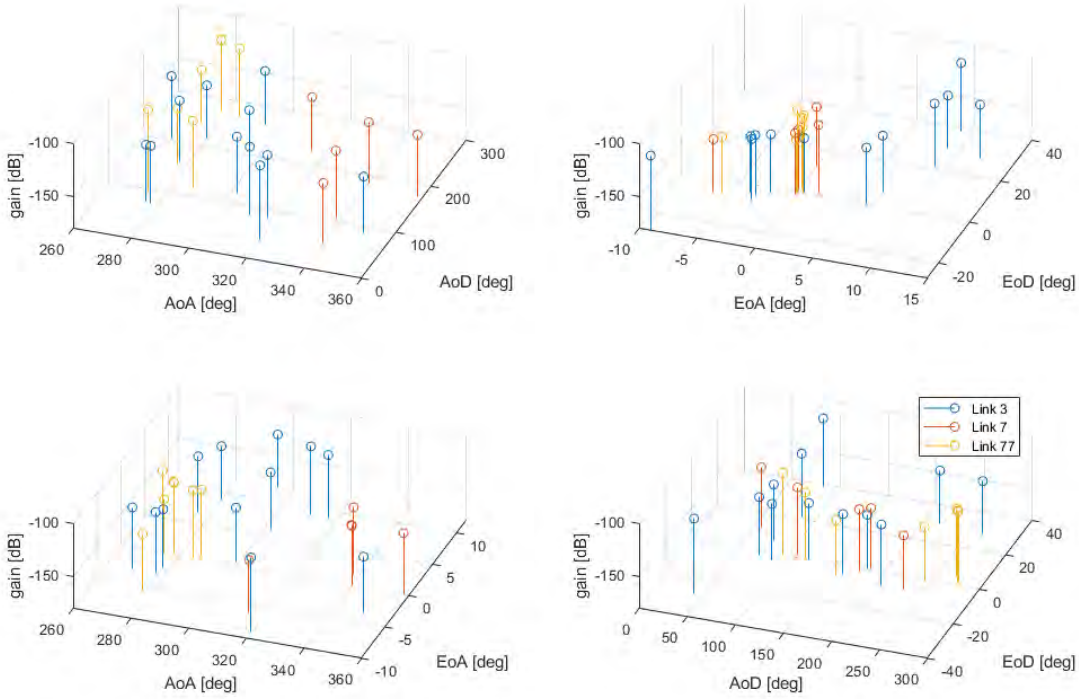


Figure 3-13: Angles and path gains of three example links of the stored channel model.

3.3 Summary

This section presented a Hexa-X channel model, consisting of 1) a list of complex permittivity estimates for materials that make up our living environments for a frequency range from 2 to 260 GHz and 2) wideband MIMO channel model at 140 GHz based on measured spatio-temporal channel sounding data in indoor and outdoor scenarios. The latter is implemented in MATLAB and is published in Hexa-X Zenodo community [DHK23].

4 Radio architecture and models

Radio architecture for the radio frequency (RF) and the first stages of digital signal processing on L1 physical (PHY) layer are implemented using customized hardware (HW) elements, and therefore, architectural work is mostly focused on the physical properties of the HW including antennas, RF circuitry, data converters and the highest-speed and the highest-resolution digital parts. Software (SW) and associated SW architecture for control is utilized for configuring and tuning different elements based on the commands from higher protocol layers and mapping these to HW. In addition, some calibration routines to improve the performance of the RF, analogue and mixed-mode HW with their appropriate timings are run on SW.

This section is highly focused on key aspects of generating functional HW architecture for RF and mixed-mode processing and means to optimize it for different target data rate and communications range requirements considering some key properties of the waveform. (Implementing SW to support those functions are considered relatively straightforward but work intensive task when implementing prototypes and products at a later phase of the development.)

Initial system modelling principles and background for calculations were given in earlier deliverable [HEX21-D21] of the Hexa-X project as in Figure 4-1. The model was refined, and detailed but highly abstracted models were extracted from different building blocks (like low noise amplifiers (LNAs), power amplifiers (PAs), analogue-to-digital converters (ADCs), etc.) in deliverable [HEX21-D22]. Some of them have been refined in this document [HEX21-D23]. Those have been used directly whenever available with short reference to the document. Scenarios modelled in the example in this section are based on Table 2-1 if not stated otherwise. Radio channel has been mostly modelled based on line-of-sight (LoS) link but specific notes are made based on the article "How many beams does sub-THz channel support?" [KGH+22]. All of the Hexa-X [HEX21-D2x] deliverables and related articles describe refined aspects and specific models for radio channel for different scenarios that can be used instead of LoS whenever needed. However, LoS presents in most cases the best possible performance that cannot be exceeded. As link range will be highly limited in most cases in Sub-THz bands this maybe at the same time the required and the only reasonable condition. However, some viewpoint on non-line-of-sight (NLoS) channels will be given.

For the same reasons as for radio channel also the HW models used in this section are quite ideal, but they try to indicate the best possible state-of-the-art (SoA) performance without implementation margin. Required link budget margins need to be evaluated in case-by-case manner and they can be several dBs or even more in some cases. Those are in some cases commented in the text at least as general level.

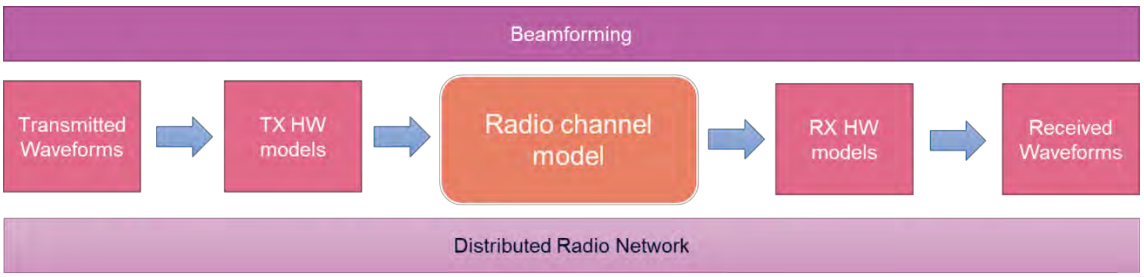


Figure 4-1: System model for 6G radio analysis from end-to-end link perspective. Beamforming, MIMO, and distributed radio network need to consider the whole signal path and limited isolation (i.e. coupling) between orthogonal channels. [HEX21-D21].

4.1 RF transceiver architecture modelling methods

What is the correct architecture for 6G radios operating at upper mmWave frequencies (100-300GHz) and can provide data rates up to 100Gbps or even above? Fully digital MIMO is naturally the preferred option from the flexibility perspective, but it suffers from high power consumption and computational

complexity especially at high data rates and bandwidths. Moreover, already in FR2 bands of 5G NR most transceivers moved to sub-array-based RF architecture, where each sub-array consisting of multiple phase and amplitude steered antennas was pointed towards beam containing one data stream (or sometimes multiple streams). With array gain and spatial filtering, the content of each stream was possible to digitize and process in data converters and digital size with reasonable sampling rate, resolution, and power consumption. With advances in technology the border frequency between fully digital and hybrid approach will gradually rise. However, the technology evolution in the area of RF and mixed-signal processing is moderate (or one could say sometimes even slow) compared to pure digital (Moore's law) and therefore when moving towards higher data rates and thus higher carrier frequencies it is anticipated that sub-array-based RF transceiver architectures are likely to be de facto also in extreme data rates in 6G.

The sub-array-based approach is sometimes called as a hybrid-MIMO architecture. However, it is only one specific (but the most straightforward and many times the most practical) case of that. The approach has been theoretically generalized in [ARA+14]. However, the generalization will lead in large antenna arrays to implementation complexity especially in RF wiring and interconnects that is not possible to realize. Only limited scale trials have been demonstrated for this architecture summing signals to multiple paths only from few antennas [MSH+18]. One must also note that in any hybrid architecture dynamic range requirements of all RF and analogue blocks change compared to fully digital MIMO. (NOTE: If multiple signals are transmitted from the same antenna power amplifiers (PAs) and any other components in the signal chain must be dimensioned taking this additional dynamic range into account [TAT+16]. The same applies for blocks in the receiver after any signal combining. All signals within the sub-array are amplified to achieve higher SNR but also at the same time larger dynamic range from the following components.)

For the reasons above, for any 6G transceivers the simplest sub-array-based hybrid architecture is the best reference for the architectural studies. The same approach reduces to fully digital approach if the number of antennas in each sub-array is only one. Naturally, in that case it makes only sense to perform phase shifting in digital domain removing any analogue phase shifter from the RF parts. The same model is also applicable for lens antenna-based solutions as the amplification comes from the lens and opportunities to further steer the beam are limited to activate the appropriate antenna feed for specific direction [AAK16]. As long as the sub-arrays are combined into one digitized stream in one or multiple phases the number of antennas is not fundamentally limited to any specific value. Of course, it has direct impact to the performance partitioning of the receiver or the transmitter impacting power consumption etc. But combining is feasible and recent implementation shows 384 element phased array consisting of multiple Radio Frequency Integrated Circuit (RFICs) and modules. Finally, the inverse relation between antenna array/element gain and beamwidth as shown for example in [RKL+20] equally valid for antenna arrays and lenses should be noted.

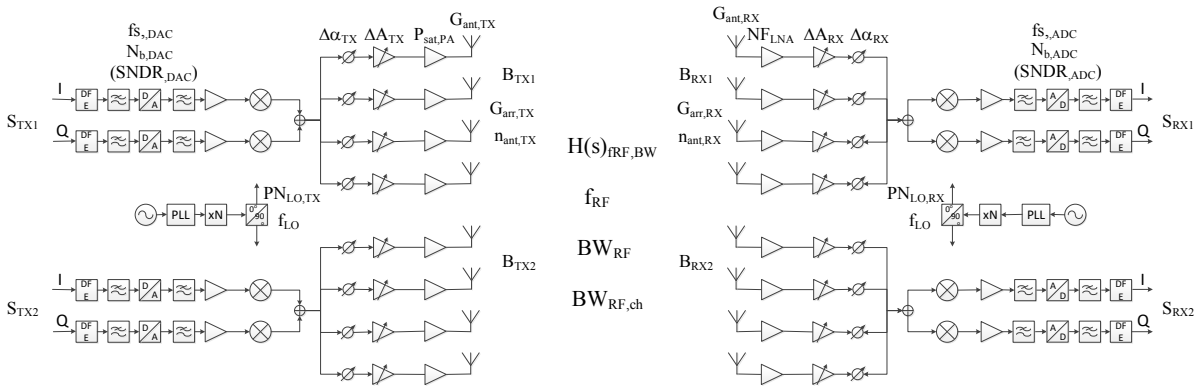


Figure 4-2: RF transceiver reference architecture for hybrid beamforming based on sub-array per beam principle. Key parameters for model abstraction are named.

A block diagram of the reference transceiver architecture for 6G analysed in this document is given in Figure 4-1. It contains key parameters that are needed in the analysis. The highest level of abstraction provides means to analyse coarsely the link performance, architectural constraints and boundaries of

various technologies using scalar performance values from different entities. The basics of the approach is described in detail in [TTP17]. The key restriction of the method is ignorance of the frequency responses in the signal chain and other components. Therefore, all impacts from frequency selective filtering, beam squint, NLOS channel, etc., should be modelled using implementation margins that are related to specific implementation or physical boundaries in the particular waveform, HW realization or both at the same time. However, this is a conventional approach when defining specification limit for example for minimum signal-to-noise ratio (SNR) for the receiver/detection. However, over the different generations of wireless systems this approach has proven to be accurate enough for architecture and coarse performance tuning of the RF transceivers.

The scalar method uses still true bandwidth estimates for the signal ($BW_{RF,ch}$), radio band (BW_{RF}) and ADC or digital-to-analogue converter (DAC) sampling rates ($f_{s,ADC}$ & $f_{s,DAC}$). In the first order, this could be understood as ideal brick-wall filters with linear phase response, possibly with some scaling. Fortunately, when designing with actual (complex) waveforms (Ex. in Matlab or other waveform-based test benches), more elaborate channel models and realistic filters, the same performance parameters used in the simplified model are applicable as it to represent HW performance in nominal, and possibly also in other environmental conditions. Of course, those would require more detailed HW simulations in the background. Just the methods to calculate signal quality using error vector magnitude (EVM) or any bit/symbol/packet level error rates requires more complex end-to-end simulations using typically baseband equivalent modelling. Many of the scalar parameters are applicable to those models as well. If more precision is needed, signal levels including waveform, noise, nonlinearity etc. could be refined with their frequency responses. However, then the accuracy of the model needs to be significantly better and for example memory effects of PAs will require additional details of the model and in most cases also the dedicated HW implemented for the function. The detailed limits of HW become then complicated and in many cases relationship to system analysis blurry. The most important aspect however is to consider the absolute power levels when operating the models in link level analysis or transmitter or receiver specifications as was done in [TTP17]. The impact of the absolute signal level is introduced for example in Section 3.2 of [HEX21-D22]. The dynamic range depends on many aspects like signal (i.e., thermal noise) bandwidth, link range, gain control both in transmitter and receiver, number of antennas and its relation to linearity.

The lack of proper modelling of signal level dependence is one of the key reasons for misunderstanding the RF constraints and opportunities. This is natural as making a detailed and accurate model requires detailed understanding of the HW implementation boundaries and their relations to HW architecture. Modelling those properly is a complex and tedious task and requires a lot of expertise. Therefore, models are often developed only to certain simulation cases with a smaller validity range. This is acceptable if the validity of the models is thoroughly understood and obeyed in the analysis.

For additional details of implementation and technology aspects reader is recommended to familiarize with SoA publications from IEEE Solid-State Circuits, Microwave Theory and Technology, and Antennas and Propagation Societies. Overviews can be found for example from [PAB+20] and [GFF+21].

4.1.1 Modelling from signals to RF performance

Also, and maybe especially, for new concepts a systematic and well-defined approach is the key to design and evaluate RF transceiver performance requirements as stand-alone entity and as a part of the end-to-end analysis in link budget or higher levels. The order to define the system can follow the following listed order (with natural feedbacks when some tuning is needed). Of course, in some cases iteration over various steps is required but even this can be supported in this flow diagram type of methodology.

1) Data rate
    a. Waveform with SNR/EVM requirements
    b. Channel BW
    c. ADC/DAC sampling rate and resolution/dynamic range
2) Link range

a. Performance partitioning between transmitter (Tx) and receiver (Rx)
    i. Downlink and uplink might be different
    ii. Here they are assumed to be symmetric
b. RF performance
    i. Rx noise
    ii. Tx output power and linearity (i.e. backoff)
    iii. Phase noise
    iv. Coarse budget for other non-idealities (image rejection, phase accuracy, drained current (DC) offset, …)
c. Waveform impacts on RF performance
    i. Required backoff /margins

3) Internal RF performance requirements
    a. Partitioning of functionalities between analogue and digital (front-end)
    b. Additional margin for data converters
    c. Input compression point (ICP) of the Rx
    d. Block/component level performance partitioning
4) RF frequency bands and spectrum
    a. Spectrum availability
    b. Channelization (frequency, MIMO, etc.)
    c. Carrier aggregation (CA) to optimize the system (HW perspective)

In order to progress through the steps case of mid-range wireless access case at 140GHz from Table 2-1 is analysed in this context. Several assumptions will be discussed and from several options only one is selected for further steps to keep the example reasonable. There will be several options in many of the steps and the selections are not necessarily present here the best or feasible choices when going towards 6G. Also, from RF technology options only Silicon Germanium (SiGe) Bipolar complementary metal oxide semiconductor (BiCMOS) is analysed here although models described in this document and elsewhere for other semiconductor technologies. SiGe BiCMOS is an appropriate reference for many lower mmWave solutions (Ex. 5G NR) and has less pronounced penalty when moving to upper mmWave region.

4.1.1.1 Data rate

First step in HW design (both for RF transceivers and digital modems) starts from maximum data rate of a specific use case, group of use cases that will be served simultaneously or total number of users and applications that are simultaneously served by the same radio protocol (or in broader sense equipment). The latter is specifically true in devices for infrastructure like base stations. Also the radio path latency requirement set constraints for the instantaneous maximum data rate. On the other hand, reliability when transmitting through radio channel makes the third important constraint leading to rate reduction due to coding and other techniques increasing the probability of successful detection. This leads to additional bandwidth requirements related to user data rate that will have more severe impact to implementation at extremely high data rates.

Although anything above is well known, the detailed analysis to define required data rate is not trivial. As latency is typically dominated by higher level protocols and hence it is not typically a problem of the RF and digital front end HW itself. Therefore, HW aspects can be mostly ignored especially in RF/analogue domain and hence the latency analysis is more related to the time-domain arrangements and waveforms that are discussed in Section 5.1.1.

Once the detailed enough data for all above are derived, instantaneous data rate can be defined. Here it is assumed that the 100 Gbps data rate based on Table 2-1 will include all higher-level protocol overheads etc. but not L1 (convolutional) coding and bandwidth overheads required for filtering in channelization (i.e. dividing to multiple frequency channels). Table 4-1 summarizes the user data, coding and frequency parameters in the example and some alternative parametrizations that could be tested

Table 4-1: Common RF and waveform parameters (assumptions) used in the example analysis.

Parameter	symbol	example	alternative
User data rate with protocol overhead	R_u	100 Gbps	1Tbps
coding rate	R_c	5/6	uncoded
carrier frequency	f_c	140 GHz	300 GHz

In this example, two basic waveforms are used i.e., single carrier QAM and OFDM with QAM subcarriers. As protocol overheads are not counted for, the numbers cannot be compared in the case of OFDM directly to achievable 5G NR data rates where overhead is quite significant.

First the bandwidth (BW) and the (data converter) sampling rate vs. data rate will be defined for various modulations. As BW and sampling rate are somewhat different, they need to be treated separately. Required RF bandwidth reserves some guard bands to enable channel selection and defines then channel raster. Generally, it can be defined for any data rate in relative scale. For single carrier (SC) modulations pulse shaping filter defines the additional BW. A typical number of roll-off r=0.35 is used here. In case of OFDM the filtering shape of the modulation is sharp but for analogue/RF filtering typical 10% guard band as in 3GPP is adopted. When calculating total spectral efficiency this must be considered. Thus, any fixed data rate can be defined for the required bandwidth vs data-rate (defined here as bandwidth relative to the user data-rate, BW/R_u) with different modulations and coding rates. For SC waveforms, the required relative BW can be written as

$$BW_R(SC) = \frac{1}{(1+r)R_c M}, \tag{4-1}$$

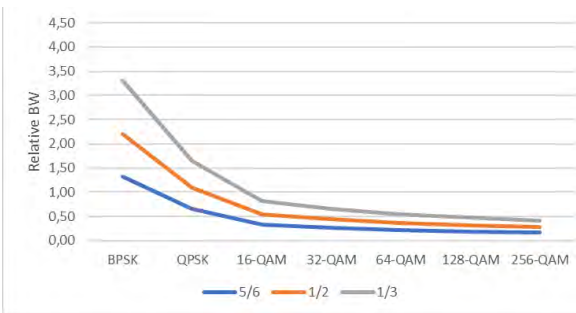
where r is the factor for raised cosine filtering, R_c coding rate and M modulation index. Similarly, for OFDM the relative BW is

$$BW_R(OFDM) = \frac{1}{(1+G)R_c M}, \tag{4-2}$$

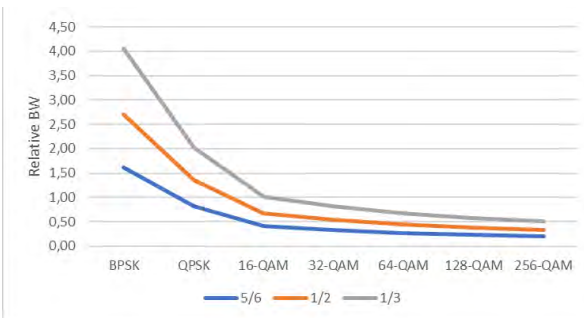
where G is the relative (nominal) guard band between the channels. Based on the values given above, the relative BWs for fixed data rate are defined in Table 4-2, whose content is illustrated in Figure 4-3.

Table 4-2: Required bandwidth for fixed data rate using different modulations with guard bands.

	M	BW_R(OFDM)			BW_R (SC)		
Code rate R_c		**5/6**	**1/2**	**1/3**	**5/6**	**1/2**	**1/3**
BPSK	1	1.32	2.20	3.30	1.62	2.70	4.05
QPSK	2	0.66	1.10	1.65	0.81	1.35	2.03
16-QAM	4	0.33	0.55	0.83	0.41	0.68	1.01
32-QAM	5	0.26	0.44	0.66	0.32	0.54	0.81
64-QAM	6	0.22	0.37	0.55	0.27	0.45	0.68
128-QAM	7	0.19	0.31	0.47	0.23	0.39	0.58
256-QAM	8	0.17	0.28	0.41	0.20	0.34	0.51



(a) OFDM.



(a) SC (roll-off r=0.35).

Figure 4-3 Required relative BWs for fixed data rate with different modulations and coding rates.

4.1.1.2 Sampling rate and resolution requirements of data converters ADC & DAC

Nyquist rate for baseband (BB) in-phase and quadrature (IQ) signals is two times the baseband BW i.e., equivalent to RF BB. To recover also clock, the phase information and sample rate twice at the RF BW (four times BB BW in both I & Q channels) are required. Therefore, ADC sampling rate is defined simply as twice the coded data rate using coding rate and modulation index as

$$f_s = \frac{2}{R_c M} R_u. \tag{4-3}$$

For 100 Gbps data rate the sampling rate (Gsps) can be defined for different modulations and coding rates according to Table 4-3. Note that the sampling rate requirement in Table 4-3 are with assumption that only the actual signal is quantized/sampled. However, in many cases the requirement is higher, for example, when using correction techniques such as digital predistortion (DPD) to correct transmitter nonlinearity. In case of transmitter linearization, so called feedback or measurement, receiver requires often higher bandwidth (1.5-3 times) [LWZ20] to provide enough bandwidth for linearization, depending on the target linearization performance.

Table 4-3: ADC sampling rate (Gsps) for 100 Gbps data rate for different modulations and coding rates.

Code rate R_c	5/6	1/2	1/3
BPSK	240	400	600
QPSK	120	200	300
16-QAM	60	100	150
32-QAM	48	80	120
64-QAM	40	66.7	100
128-QAM	34.3	57.1	85.7
256-QAM	30	50	75

Moreover, it would be desirable to use the same ADC resources for both observation receiver and actual data receiver, which expands the sampling requirements [Ter22]. DAC sampling rates can be defined equivalently. Note that similar bandwidth extension is needed also in the entire RF signal path. This is because the predistorted signal bandwidth is larger than the actual signal bandwidth before the PA.

Table 4-4: ADC dynamic range and sampling rate requirements for various combinations resulting to 100 Gbps data rate with 5/6 coding rate.

OFDM					SC (r =0,35)				
	RF BW factor	**BW GHz**	**fs GHz**	**SNDR ADC**		**RF BW factor**	**BW GHz**	**fs GHz**	**SNDR ADC**
16-QAM	0.33	33	60	44.6	**16-QAM**	0.41	40.5	60	41.7
64-QAM	0.22	22	40	48.4	**64-QAM**	0.27	27	40	45.7
256-QAM	0.17	16.5	30	55.7	**256-QAM**	0.20	20.25	30	53.1

To evaluate ADC requirements and power consumption also required resolution for different waveforms needs to be evaluated based on their minimum SNR requirements also considering some margin for clipping and the fact that ADC cannot dominate the total noise figure (NF) of the receiver and is preferably almost negligible. Therefore, ADC quantization noise must be roughly 15 dB below the amplifier input referred noise from RF and analogue components giving only 0.1dB penalty to input referred noise level. However, this is one factor that significantly increases ADC resolution requirement compared to classical digital only noise estimates for digital modulations. Also, some margin needs to be reserved for gain control etc. Details of these analysis are given in public deliverable [HEX21-D12]. Based on that resolution and sampling rate for different modulations are defined for small coding rate or 5/6 i.e., some coding will be used but as little as possible to minimize bandwidth overhead. The results for selected modulations are given in Table 4-4 and illustrated in Figure 4-4. It is assumed that

transmitter and receiver split the EVM budget equally as is quite common practice in very high data rate scenarios in 5G NR.

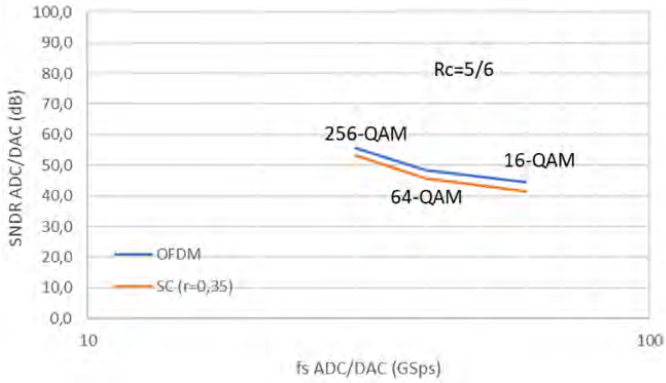


Figure 4-4: ADC dynamic range and sampling rate requirements for various combinations resulting to 100Gbps data rate.

4.1.1.3 ADC analysis based on the requirements

Next ADC requirements have been added to chart 3-14 from [HEX21-D12] that is based on data by ADC comparison from Boris Murmann [Mur22]. The data is checked against the recent version as referenced. In Figure 4-5, targets marked as solid lines have been placed to the data plotted against equivalent Nyquist sampling rate of the reported converters. In order to evaluate feasibility of the ADCs, two trend lines as a function sampling rate have been drawn based on the best SoA signal-to-noise and distortion ratio (SNDR) achieved for any frequency. One can see that the commercial converters are well below the limits while the anticipated 6G ADCs would still go above the feasibility. Extensive research is on-going to achieve the targets in data converter community. However, achieving both good dynamic range and high sampling rate is not obvious at those frequency ranges. Therefore, solutions in system design level require also serious considerations of the waveform. Those include at least i) reducing the sampling rate with more parallel processing i.e. carrier aggregation, or ii) reducing crest factor of the waveform to reduce dynamic range of the converter while increasing inevitably the bandwidth and thus sampling rate as below. Higher bandwidth and lower resolution are closer to the realistic implementation according to trendlines drawn to Figure 4-5.

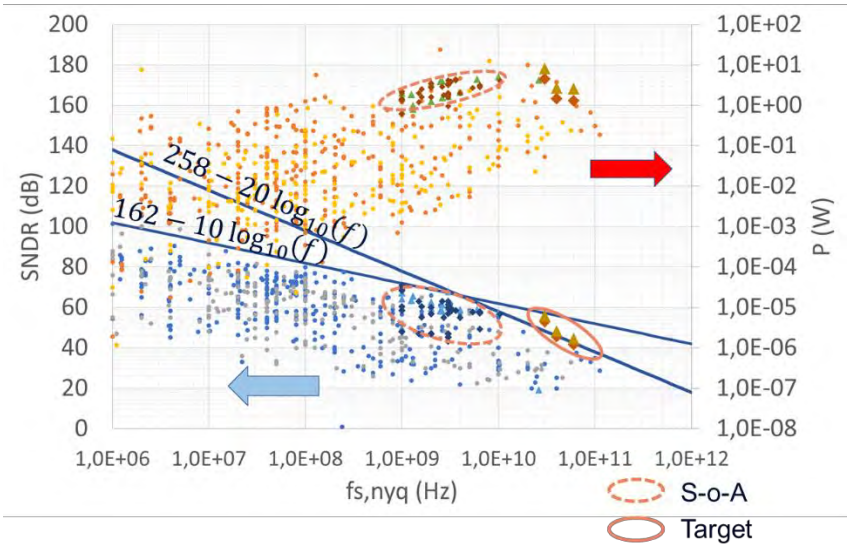


Figure 4-5: ADC dynamic range (SNDR) and power consumption (P) is compared to ADC dynamic range requirements for the OFDM and CW waveforms (solid oval). Some commercial state-of-the-art ADCs with high resolution and sampling rate (dashed ovals) have been added to complement the data from [Mur22].

ADC models based on data in [Mur22] with selected commercial SoA converters can be further refined based on the formulas presented by Walden [WAL99] and Schreier [ST05] with boundaries defined in [Mur22]. However, drawing conclusion simply by using the formulas is not straightforward and dependencies between dynamic range and power consumption cannot be drawn just based on boundaries. Data in [Mur22] is captured from various scientific publications that are prototypes to show feasibility or new ideas and they may not necessarily include all the components comprehensively like clocking needed in complete ADC sub-systems. Therefore, comparison requires also perspectives to data available publicly from commercial ADCs. It is well-known that ADCs are somewhat more complex circuits compared to corresponding DACs having similar sampling rate and resolution. Therefore, DACs are known to consume less power and if ADC is feasible, it should be possible to have also corresponding DAC available. As DAC power consumption and performance depends also on (analogue) load impedance it is driving. Therefore, it is more difficult to make similar predictions of figure-of-merits (FoMs) for DACs as presented for ADCs in [Mur22].

4.1.1.4 Output power vs. Psat (backoff) and waveform

Defining nonlinearity or limits for it is far from being straightforward especially in case of PAs. That is also the same for a receiver where for the received signal a typical measure is ICP1, which is typically defined as 1 dB compression point from linear gain. As in receiver, the linearity is mostly rather smooth and follows 3rd-order nonlinearity, it is possible using this model to define EVM as a function of input backoff (BO) from ICP1 dominated by 3rd order term. The simulated behaviour is shown in Figure 4-6 for OFDM based 5GNR signal and for single carrier (SC) 64-QAM modulation without any noise or other nonidealities. SC modulations from 16QAM to 256QAM show slightly different behaviour to each other while OFDM has almost the same envelope independent of the sub-carrier modulation. Difference between OFDM and SC is clearly visible in Figure 4-6 i.e. SC shows better performance close to compression due to lower crest factor i.e. peak-to-average rate (PAR) of the signal. This model can be used especially when evaluating receiver non-linearity and its performance requirements against the required dynamic range.

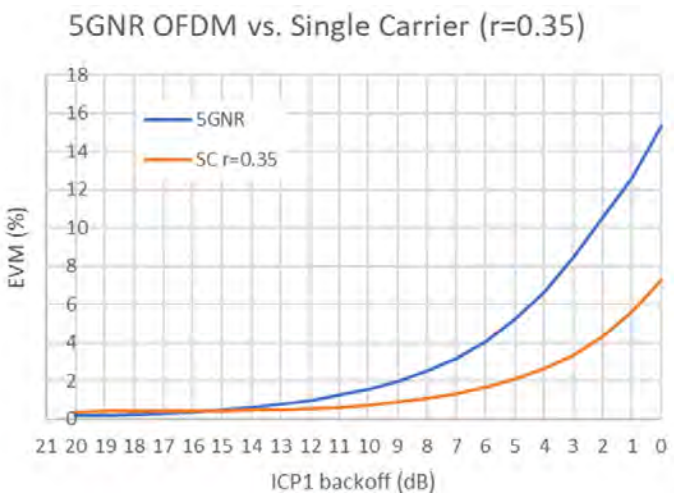
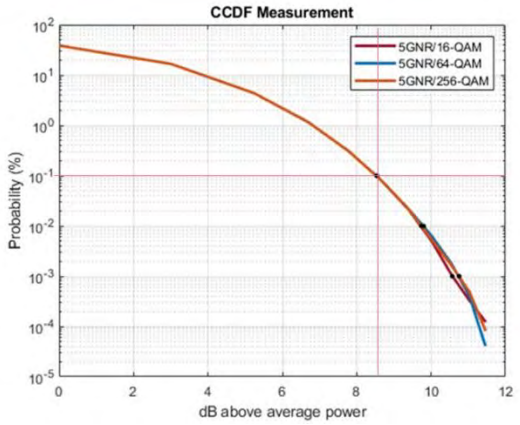
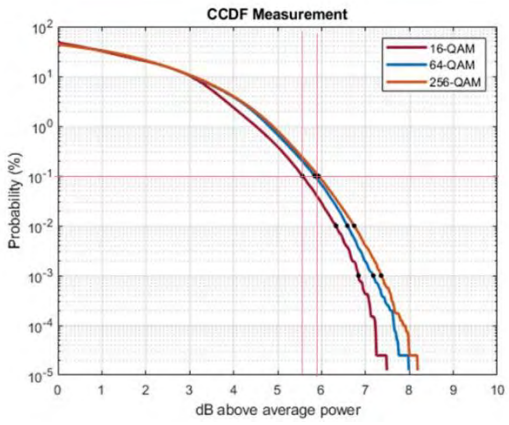


Figure 4-6: EVM for 5GNR and SC (r=0.35) 64QAM modulations from the 1 dB compression point of the receiver (or transmitter) where non-linearity is dominated by the 3rd order term.

For a transmitter, definition of nonlinearity requirements depends not only on the crest factor but more complex non-linear behaviour (depending on the PA class, technology and design aspects), opportunity to clip the highest peaks somewhat even in digital side reducing PAR, and possible linearization. Those have direct impact how close to saturate output power (P_{sat}) one can drive the PA. In order to make first-order estimate on the relation of P_{sat} to a class-A type of PA without linearization or digital clipping one can estimate necessary backoff from probability of the peaks. In this case, BO for different modulations is determined based on 0.1% peak probability as shown in Figure 4-7 leading to 5.6-8.5 dB BO depending on the modulation. This is the major factor for the link range when evaluating the PA output power for the same PA. Values are summarized in 4.1.1.5.



5G NR modulation



SC, r=0.35 modulation

Figure 4-7: Probability of the peaks above average signal power.

4.1.1.5 Minimum SNR/EVM for radio link

Finally, minimum SNR/EVM budget over the whole link for different modulations need to be divided between corresponding transmit and receive signals budgets to evaluate the link range. In this case EVM budget has been equally shared between transmitter and receiver that is a typical way in 3GPP standards for high-speed signals. Also, the requirements are taken from 3GPP specification whenever available. Results are summarized in Table 4-5. These are the values to be used in the link analysis in Section 4.1.2.

Table 4-5: Minimum SNR/EVM and PAPR (i.e. required BO in this case) for different modulations in the analysis.

	modulation order	3GPP 5G NR SNR min TX+RX (dB)	3GPP 5G NR SNR min TX+RX (%)	3GPP SNR min TX/RX (dB) equal split	3GPP SNR min TX/RX (%) equal split	PAPR SC r=0.35 no clipping	PAPR OFDM 5GNR no clipping
QPSK	2	12.1	24.7 %	15.1	17.5 %		
16-QAM	4	15.1	17.7 %	18.1	12.5 %	5.6	8.5
64-QAM	6	18.9	11.3 %	21.9	8.0 %	5.8	8.5
256-QAM	8	26.2	4.9 %	29.2	3.5 %	5.9	8.5

4.1.2 Link range analysis using 100 Gbps as an example

RF link analysis can now use various LNA noise (representing whole receiver noise) and PA saturated power values with modulation specific backoff to represent margins to avoid signal quality deterioration as described in Section 4.1.1. Before analysis some comments on the validity of the analysis and key parameters will be given

4.1.2.1 General considerations on the use of the model

The following aspects need to be considered and understood when interpreting the results:

- Transmitter power is based only on power amplifier models.
- Receiver NF is based only on the model of LNA and not on the whole receiver NF, which is typically somewhat higher than the stand-alone LNA.
- Coarse, experience-based estimate on the antenna element gain (directivity and efficiency) and receiver noise.

- Front-end losses between antenna elements and active circuitry (PA & LNA) have been neglected ('zero loss').
- Tx and Rx are assumed to have separate antenna arrays i.e. no transmit-receiver switch (TRx switch) is included that has inevitable loss both in transmit and receive directions.
- Analysis is based on scalar values and flat filtering responses that are not feasible in real implementations. Possible effects of physical bandpass filters and resonators and beam squint due to phase shifter in typical arrays have been ignored.
- Perfect beam alignment is assumed (no beam pointing error estimated).
- Only LoS channel model is adopted here but more sophisticated models can be applied later.
- Non-linearity aspects would require more elaborate models and analysis both in transmit and receive ends of the link.
- Only single-carrier link is analysed here but mechanisms to enhance this to multi-carrier / carrier aggregation schemes is quite straightforward.
- Power consumption and therefore also heat, heat management, and form factor may and likely will be a limiting factor in cases when a large number of antennas is required. It is possible to enhance the analysis to that direction but it is not done yet.
- This analysis gives guidelines for more detailed power consumption estimates but details are not given in this section. Some modelling aspects with a concrete example have been given in [HEX21-D21] and in [HEX21-D22].
- Models are specifically done for 100-300 GHz range but they give also guidelines over a broader frequency range. However, especially, frequencies below 10 GHz should not be directly compared using the models before further characterization of their validity at those frequencies.
- Model validity needs to be understood and acknowledged.

Therefore, all results given in the following can be considered as **fundamental technological boundaries that cannot be exceeded** but definitely there will **be additional sources of loss** either deteriorating the communications range or causing a need to add parallelism (i.e. number of antenna elements) to compensate the impact. Analysis is made for phased arrays but can be quite straightforwardly enhanced to other antenna types and solutions.

Despite of all disclaimers as above **the model and results are valid in the relative scale and they give also a coarse ballpark for absolute values** that is a core requirement for any implementation and architectural estimate to get start with. The proposed method gives opportunities to narrow the gap between conventional communications system design and HW architecture and implementation design process at an early stage. However, it will not remove the need to mutually optimize and refine the system based on the same principles when standardization and product development are considered.

The proposed model and the method give means to:

- model system boundaries over a large range in RF spectrum based on the data rate requirement.
- model with key parameters including noise and output power over broad range of radio spectrum and get guidelines for preferred range of spectrum for particular application with certain data rate requirement.
- utilize highly systematic and solid approach over large frequency range.
- enhance modelling to different sub-topics and bring more accurate models for analysis.
- give concrete inputs to model system with more accurate models in waveform simulations (Ex. MATLAB), RF/mixed-mode hardware simulations (Ex. Cadence, ADS, CST, HFSS) and digital hardware simulations (Ex. System-C, Verilog, VHDL).
- bring abstracted results of the refined simulations back to link level analysis from all domains involved.
- co-design system at different hierarchical levels.
- facilitate and maintain consistent understanding of the key aspects need to be understood from the system.

The results in the following examples are compared to each other using the required number of antenna elements in a phased array into account. It assumes that PA output power, Rx NF or any other parameter will not change when parallelism in HW is added/enhanced. That is not generally valid and further refinements are needed based on the studies on antenna isolation, its impact to power amplifiers, examples of complete phased array transceiver implementations etc. are somewhat included. However, such examples are always based mostly on specific scenario/HW involved to analysis and their generalization is not easy. Therefore, those can and should be considered only when architectural design on a specific HW platform will be started. One just needs to understand that some implementation margin needs to be accepted at a later phase.

4.1.2.2 Modelling example for 100Gbps communications

Parameters of the analysis are listed in Table 4-6. The analysis is done by sweeping the frequency from 28GHz (as 5G NR reference) up to 300GHz. In order to keep analysis and specifically number of results reasonable only two range options (10 and 300 meters) for communications range, two modulations (16/64-QAM) and single carrier (SC) vs. OFDM are compared. The example is done here only for one semiconductor technology (SiGe) but similar analysis is possible also for other technologies when modelled similarly. Those are partially described in this document and in some models in [D2.2]. Here it is assumed that the longest link range is only determined by transmitted power (including PA backoff) and receiver noise. Phase noise with detailed characterizations in this document is still ignored although it will have some impact on many aspects described here. However, it is necessary to first see the impact of more conventional limits and then complement the analysis sometimes later with phase noise impacts also to range. Ideas of the requirements and impact are described in detail elsewhere in this document.

Table 4-6: HW models used in the analysis.

Parameter	symbol	example
User data rate with protocol overhead	Ru	100 Gbps
coding rate	rc	5/6
carrier frequency	fc	28…300 GHz
Target range	m	300/10
Psat PA model (SiGe) (Figure 4-15)	dBm	$P_{sat}(f_c) = -0.097834 f_c + 33.8299$
PA BO	dB	Table 4-5
NF model (SiGe)	dB	$NF_{min} = 1.75 \exp\left(\frac{f_c}{130.7}\right)$
Antenna element gain	dBi	5
number of antennas		To be calculated
modulation		16/64-QAM
BW-RF	GHz	Based on modulation and data rate
Radio channel		LOS and -10/-20 additive attenuation (NLOS)
Phase noise	dBc/Hz	Not included to the example below
RX ICP1	dBm	Not included to the example below

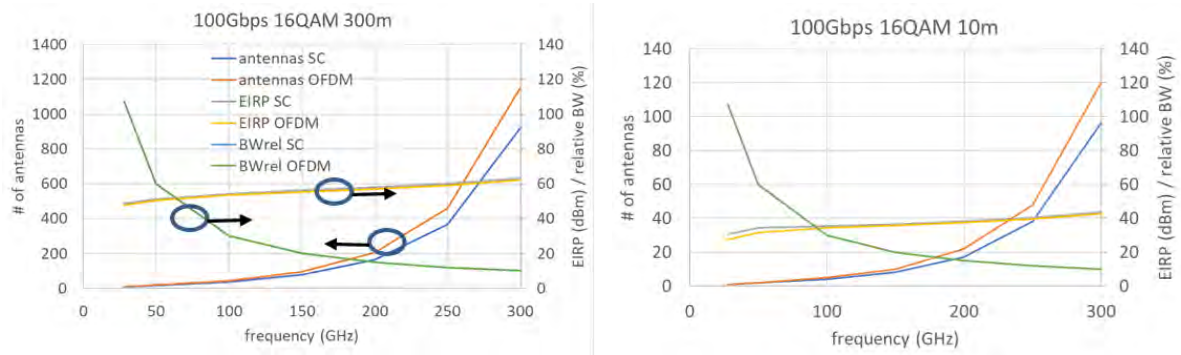
Analysis is done based on the link budget model with the conditions as in Table 4-6. Charts shown in Figure 4-8 and Figure 4-9 show required number of antennas in left y-axis, and effective isotropic radiated power (EIRP) and relative BW in right y-axis. The two latter ones are practical means to follow some practical boundaries. If relative BW exceeds 10%, HW implementation becomes more complex and exceeding 20% should be avoided by all means for a single radio channel passed through the RF front-end. On the other hand, EIRP is a regulated part of the spectrum usage and at least 75 dBm boundary should not be exceeded. In most cases this is possible. However, as it can be observed from the following charts, there are cases where the EIRP limitation becomes critical (in case the product design can be made considering the thermal aspects to meet with other constraints).

Note that the details of the legends and the corresponding vertical axis to which each curve points are shown on Figure 4-8 (a). Reader should understand that OFDM due to larger backoff requirements will require more antennas. Also, larger SNR requirement of 64-QAM compared to 16-QAM leads to larger antenna arrays in all cases although the relative bandwidth is smaller. These are key aspects when evaluating the appropriate waveform properties for 6G. The typical roughly 10% difference between

OFDM and SC modulations for the same bandwidth might sound small but that is an indicative number for differences on RF power consumption including PAs. As energy efficiency is critical now and it will be even more for 6G this is a significant difference. Saving 10% by some other means would be very difficult if not possible in apples-to-apples analysis.

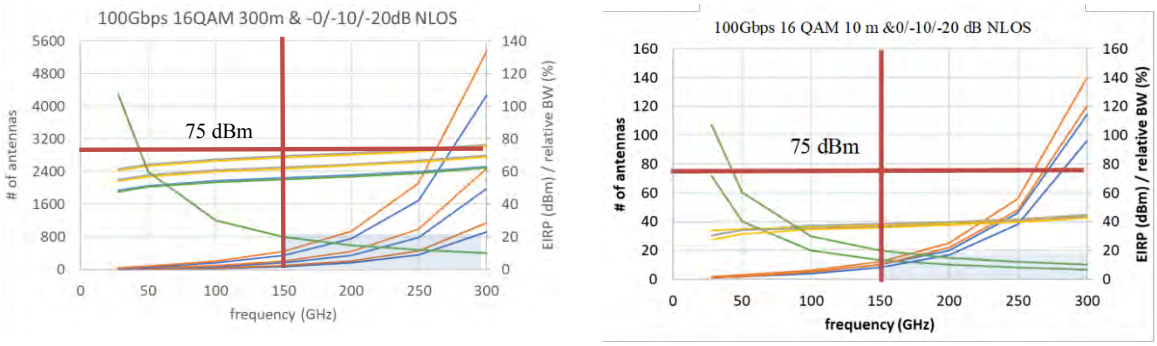
Detailed conclusions will be left for the future work, but charts given here with explanations in the charts themselves and figure captions should give good idea on some of the core aspects here.

The number of antennas, as shown in Figure 4-9, naturally needs to be increased as loss is increasing. Light blue box indicates the range when relative bandwidth is 20% or below. Also, 75dBm line to show EIRP limit is shown. It will not be a problem in this case.



(a) 16QAM, 300 m distance **(b) 16QAM, 10 m communication distance**

Figure 4-8: Number of antennas, EIRP and relative bandwidth as a function of frequency for 100 Gbps target data rate with different modulation orders and link distance. Legends and axis to which each of the curves points are shown in figure (a).



(a) 16QAM, 300m distance **(b) 16QAM, 10m distance**

Figure 4-9: Number of antennas, EIRP and relative bandwidth as a function of frequency for 100 Gbps considering 10 dB and 20 dB to describe pathloss of NLOS radio channel.

4.2 Hardware models

Nonideal behaviour of RF hardware set boundaries and limitations for the system performance. Visioned very high carrier frequencies, and wide signal bandwidth creates also additional challenges for the RF design. Key RF nonidealities that reduces the signal quality are quantization noise of the converters, IQ-imbalance, phase noise, thermal noise, and PA nonlinearity. Quantization noise and thermal noise (noise figure model) are already described in [HEX21-D22], amended by analysis given in section 4.1 of this document considering the dynamic range requirements of the data converters. IQ-imbalance as such is typically fixed and can be calibrated. Hence, here the focus is selected on the phase noise and PA nonlinearity, that are the dominant RF nonidealities. The new significant thing related to phase noise is to understand the fast phase variations that has impact due to the very high sampling rates (and symbol rates). In terms of PAs, the analysis is divided into three parts that model saturated output power, memoryless nonlinearity of the PA, and the potential memory effects. Some experimental data is also shown to collect data for the models, as well as to verify the modelling approach. Note that the definition of the hardware model in this section is not only modelling the behaviour, but also the system level performance in terms of effective signal to noise ratio for the phase noise, as well as realizable output power from the power amplifiers. The performance modelling is important, for example, in link budget calculations.

4.2.1 Phase noise

Phase noise is identified as one of the key nonidealities that limit the performance of wideband 6G systems operating in high carrier frequencies [DPS20, CHK+17]. In this section, the modelling and analysis is divided into two parts; the first one determines a reference parametrization and model for the phase noise at different operating frequencies, while the second subsection analyse the impact of the proposed model for the system performance with different centre frequencies and bandwidths.

4.2.1.1 Millimetre-wave and sub-THz phase noise for frequency-multiplied LO chains

Typical method of generating local oscillator (LO) signal for high carrier frequency is to generate some lower frequency signal with phase locked loop (PLL) and use frequency multipliers to further increase the frequency towards the desired operating frequency [HEX21-D22]. When using multiple beams (subarrays), typical approach is to route the same LO to all independent subarrays after the PLL, meaning that all subarrays share the same PLL. Since the PLL dominates the phase noise chain, the contribution of the frequency multipliers to the overall phase noise is typically not considered. From the beamforming perspective, this is good as the subarrays are then well-synchronized with respect to each other and the phase does not drift significantly over the subarrays that could cause additional challenges to the multi-beam systems especially if some interference-cancellation or zero-forcing techniques are applied. Hence, the focus is on describing the phase noise of a single LO chain.

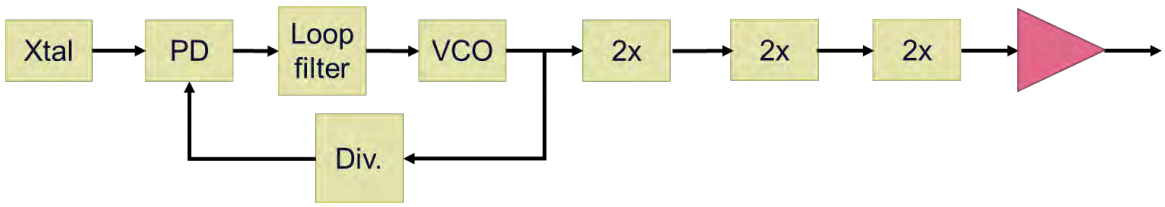


Figure 4-10: LO chain with frequency multipliers.

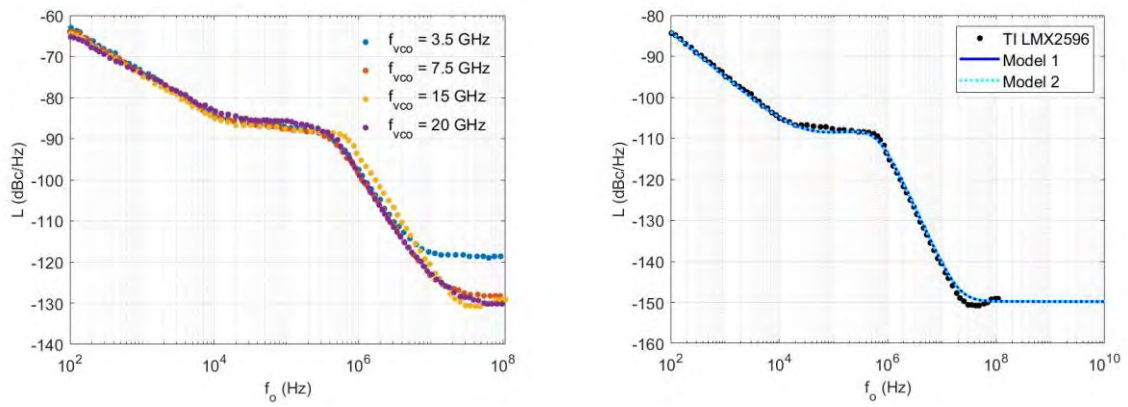
In [HEX21-D22], phase noise is modelled by the filtered gaussian phase noise approach, where the aim is to model the equivalent power spectral density of the phase noise and the actual time domain model used in simulations is, hence, an inverse Fourier transform of the model, with the assumption that the spectrum is conjugate symmetric. Similar approach is also used in 3GPP standardization with various parametrizations. In this document, two equivalent derivations for the modelling the single-sideband (SSB) phase noise spectrum are presented. The first can be written as

$$S(f_o) = 10^{\frac{N_{ref}}{10}} \left(\frac{f_c}{f_{pll}}\right)^2 \left(\frac{\prod_{n=1}^{N}\left(1+\left(\frac{f_o}{f_{z,n}}\right)^{\alpha_{z,n}}\right)}{\prod_{m=1}^{M}\left(1+\left(\frac{f_o}{f_{p,m}}\right)^{\alpha_{p,m}}\right)}\right), \quad \textbf{(4-4)}$$

where f_o is the offset frequency from the carrier frequency f_c, N_{ref} is the reference phase noise level at PLL operating frequency f_{pll}, $f_{z,1}$ .. $f_{z,N}$ are the corner frequencies of the zeros, $f_{p,1}$ .. $f_{p,M}$ the pole frequencies, $\alpha_{z,1}$ .. $\alpha_{z,N}$ are the orders of the zeros, and $\alpha_{p,1}$ .. $\alpha_{p,M}$ the orders of the poles, respectively. The frequency multiplication part is described by term $\left(\frac{f_c}{f_{pll}}\right)^2$ that shifts the whole phase noise curve up 20 dB / decade. In the expression (4-4), the flat part of the phase noise is done by the last zero, that turns the phase noise spectra into flat. Because the flat part is the most important phase noise contributor in very wideband systems, that 6G systems generally are, (4-4) can be also rewritten in another form as

$$S(f_o) = \left(\frac{f_c}{f_{pll}}\right)^2 \left(10^{\frac{N_{ref}}{10}} \frac{\prod_{n=1}^{N}\left(1+\left(\frac{f_o}{f_{z,n}}\right)^{\alpha_{z,n}}\right)}{\prod_{m=1}^{M}\left(1+\left(\frac{f_o}{f_{p,m}}\right)^{\alpha_{p,m}}\right)} + 10^{\frac{N_{th,pll}}{10}}\right), \quad \textbf{(4-5)}$$

where the additional term $N_{th,pll}$ describes the (white) noise level at the PLL output. Using (4-5) is more systematic as it helps to derive noise calculations for the LO chain with potential buffer amplifiers and frequency multipliers. It also gives easy parameter for analysing the phase noise performance with very wideband systems as single parameter tells the noise level that is then scaled up by 20 dB/decade (memory-dependent filter parts from equation (4-5) can be then neglected with analysing maximum achievable performance with perfect cancellation of the memory-dependent phase noise terms). Also, (4-5) needs generally less parameters as the flat part are generated with one additive term instead of separate zero expression.



(a) the example TI LMX2596 PLL with different PLL output frequencies all scaled to 150 GHz and

(b) the PLL phase noise spectra at 15 GHz compared to the derived models 1 and 2.

Figure 4-11: Phase noise spectra.

As an example, TI LMX2596 [TLM19] SoA frequency synthesizer operating up to 20 GHz is used. In order to find the best PLL base frequency to multiply up by the frequency multiplier, a simple comparison example can be made: Figure 4-11 (a) illustrate the phase noise spectra of the LMX2596 extracted from the datasheet with different PLL output frequencies and multiplied all to 150 GHz carrier frequency for fair comparison. As one can observe, the spectrums look very similar. The main difference is in the 3.5 GHz PLL output frequency, when the flat phase noise part seems to be raised compared to others. Otherwise, the phase noise spectra look very similar with respect to each other and especially the flat part is almost equal among the rest of the curves. From these, 15 GHz as the PLL base frequency was selected.

An example parametrization for both described models is given as in Table 4-7 and Table 4-8 and the corresponding curves at 15 GHz in Figure 4-11 (b). As one can see, both models give the same curve and their behaviour approximate also well the original PLL data given in the datasheet.

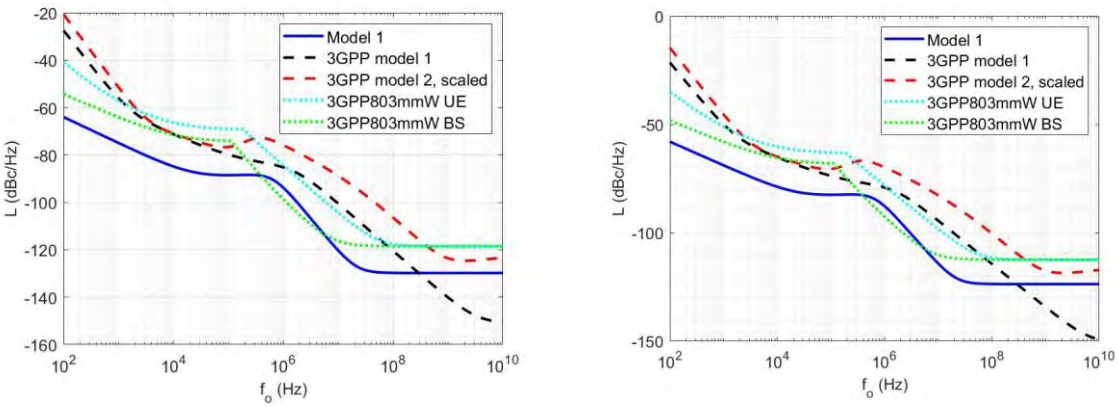
Table 4-7: Model parametrization for phase noise model 1 given in (4-4).

f_{vco}	15 GHz			
N_{ref}	−62 dB			
n, m	$f_{z,n}$	$\alpha_{z,n}$	$f_{p,m}$	$\alpha_{p,m}$
1	25 k	1.3	1	1.1
2	21 M	2.8	650 k	3

Table 4-8: Model parametrization for phase noise model 2 given in (4-4).

f_{vco}	15 GHz			
N_{ref}	−62 dB			
$N_{th,pll}$	−149.8 dB			
n, m	$f_{z,n}$	$\alpha_{z,n}$	$f_{p,m}$	$\alpha_{p,m}$
1	25 k	1.3	1	1.1
2	N/A	N/A	650 k	3

For fair comparison, the comparison among between the proposed model parametrization to the existing upscaled 3GPP reference models is illustrated in Figure 4-12 (a) 150 GHz and (b) 300 GHz. As one can observe, different models have rather significant differences in the overall characteristics.



(a) Scaled to 150 GHz **(b) Scaled to 300 GHz**

Figure 4-12: Comparison of the derived phase noise models against 3GPP. The models 1 and 2 given in this document gives the equivalent phase noise spectra.

4.2.1.2 System phase noise behaviour and limitations on the achievable performance and bandwidth

Figure 4-13 illustrate the SSB phase noise spectra together with different system level aspects that have impact on how different regions matter on the signal phase variations from the system perspective. Note that the spectrum is given in logarithmic frequency axis, as typically made, to highlight different regions. Waveform selections and frame structure has impact on the overall performance of the system under phase noise. In particular, the phase noise impact on different waveforms depends on how the phase is synchronized in the receiver and how the data and pilot sequences are distributed across the time domain signal with respect to the phase jitter characteristics.

In multi-carrier waveforms, the lower offset frequency components of the phase noise cause so-called common phase error (CPE) due to long symbols that can be compensated in the receiver by proper phase noise tracking reference signals. The higher offset frequencies and the flat part of the phase noise then causes inter-symbol interference that generally sets limitation on the achievable SNR. In single carrier waveforms, the symbols are short in time domain which means that the potential phase drift over a single symbol is mostly dominated by the flat phase noise if fast symbol rate is assumed. The potential impact of the lower offset frequency components of the phase noise depends also on the signal length,

i.e., the speed of the phase synchronization in the receiver. These two (minimum frequencies that matters for the system and the maximum frequency, i.e., symbol rate), which are defined as f_1 and f_2 for calculating the overall jitter from the phase noise in [HEX21-D22, Eq. (3-14)]. In addition to waveform level parameters, also the environment has impact on the phase variations of the received signal. For example, mechanical vibration causes random phase variations for the link in terms of Doppler effect, but also part of the vibration may be conducted to the synthesizer itself and cause phase to vary [HNH+07, HNH09]. These may be important especially in industrial applications, for example in factory automation, drones, etc. The impact of the mechanical vibration on the phase noise spectra is generally experienced in lower offset frequencies, but in the very high centre frequencies, their harmonic impacts may be seen in even tens of MHz offset frequencies.

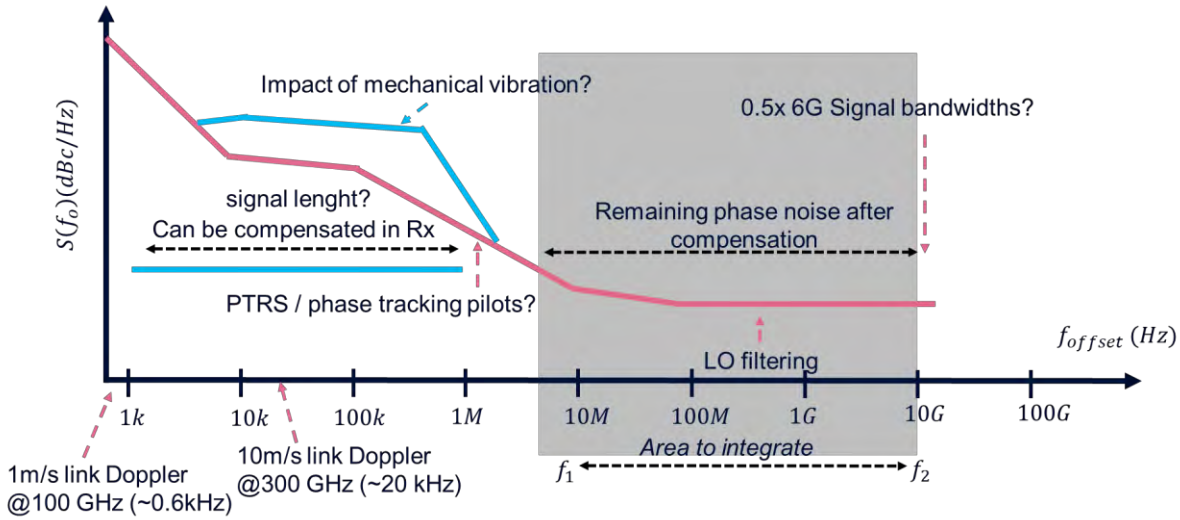


Figure 4-13: Illustration on the different aspects in system level that have impact on the importance of different regions on the phase noise characteristics in different offset frequencies.

In the literature [DPS20, CHK+17], phase noise with high offset frequencies (wideband phase noise) is identified as one of the most critical performance bottlenecks especially with very wide signal bandwidths versioned in 6G. However, the impact of the operating frequency has not been analysed in detail in the open literature in a systematic way. In the following, the limitation of the phase noise on the the achievable bandwidth and SNR in different centre frequencies is analysed. The analysis assumes that the flat part of the phase noise dominates. This is the case if the signal structure is generated in such a way that the occurrence frequency of the phase noise tracking signals is more than the corner frequency where the phase noise spectra turn to flat. In practice, this should be at least tens of MHz that is certainly possible at least with very wideband single-carrier waveforms. This basically means that when $f_{ptrs} > 2f_{z,N}$ in model 1(4-4) is chosen, the memory-dependent part of the phase noise is assumed to be compensated. In model 2 (4-5) this then means that the remaining part of the phase noise is only the flat term of the phase noise. By doing this, the remaining standard deviation of the phase noise can be written as

$$\sigma_{pn} = \sqrt{B \cdot \left(\frac{\mathrm{f_c}}{\mathrm{f}_{pll}}\right)^2 10^{\frac{N_{th,pll}}{10}}}, \tag{4-6}$$

where B is the bandwidth of the signal (symbol rate). Thus, the phase noise-limited SNR can be written as

$$\begin{aligned} SNR_{PN} &= -20\log_{10}\left(\sqrt{B \cdot \left(\frac{\mathrm{f_c}}{\mathrm{f}_{pll}}\right)^2 10^{\frac{N_{th,pll}}{10}}}\right) \\ &= -N_{th,pll} - 10\log_{10}(B) - 20\log_{10}(f_c) + 20\log_{10}(f_{pll}). \end{aligned} \tag{4-7}$$

Note that SNR_{PN} in the equation mimics the SNR budget given only for the phase noise. In addition to this, the overall SNR is naturally also limited by the rest of the RF nonidealities that set their own limits. In the end this means that the achieved overall SNR is always less than the SNR budget given for the phase noise. The equation can be also defined to calculate maximum achievable bandwidth for a given centre frequency and SNR budget as

$$B = \left(\frac{f_{vco}}{f_c}\right)^2 10^{\left(\frac{-N_{th,pll}-SNR_{PN}}{10}\right)}. \quad \textbf{(4-8)}$$

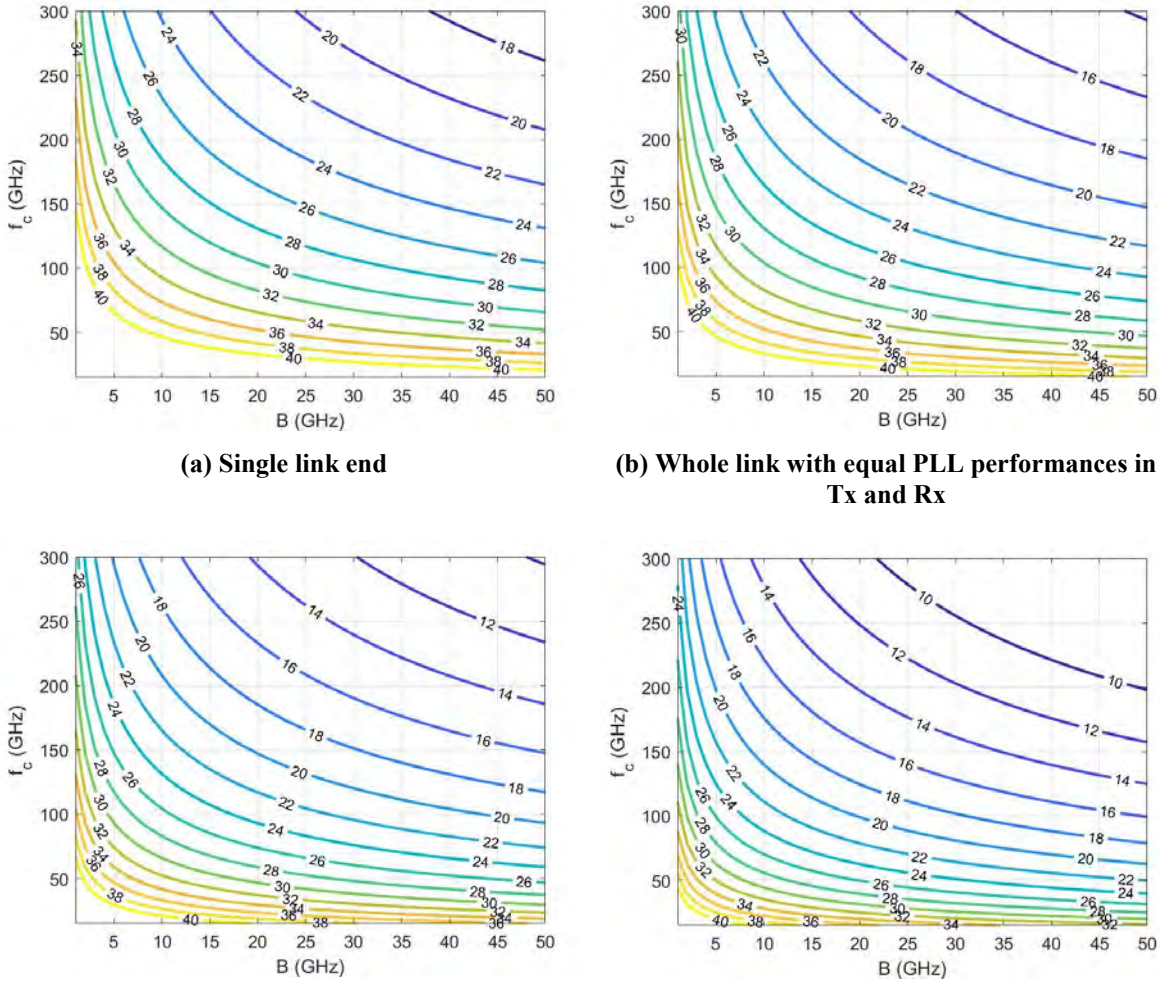
An important note from the expressions (4-7) and (4-8) is that the bandwidth dependency is 10 dB/decade, and the carrier frequency impact is 20 dB/decade. This generally means that it is doubling the centre frequency for a given fixed SNR target, decreases the achievable bandwidth to be four times less. As these both (bandwidth and centre frequency) increases when going towards 6G systems, this result is of a particular importance.

A single radio link generally includes at least a contribution from two LOs, i.e., phase noise from transmitter and receiver. Usually, in cellular links, it is often assumed that the mobile terminal has worse performance. The overall link phase noise can be calculated as the root mean square (RMS) sum of the noise contributors (Tx and Rx) in the best possible case if receiver own thermal noise and multipath propagation would not limit the performance of the phase noise compensation. To get the values of the previous analysis into context, four downlink scenarios in different carrier frequencies and bandwidth are analysed. The scenarios with the parameters are given in Table 4-9.

Table 4-9: Simulation parameters for link phase noise analysis.

f_{pll}	15 GHz		
Case	**description**	Tx $N_{th,pll}$	Rx $N_{th,pll}$
1	Only Tx contribution	-149.8 dBc/Hz	N/A
2	Symmetric link with identical PLLs	-149.8 dBc/Hz	-149.8 dBc/Hz
3	Rx with 6 dB worse PLL	-149.8 dBc/Hz	-143.8 dBc/Hz
4	Rx with 10 dB worse PLL	-149.8 dBc/Hz	-139.8 dBc/Hz

The results in all four scenarios are illustrated as contour plot in Figure 4-14 (a)-(c). The results are illustrated as contour-plots where each curve in the figure describes the bandwidth-carrier frequency pair for a given phase noise limited SNR. One should note that in lower frequency communication systems with less than 1 GHz of signal bandwidth, the remaining phase noise is rarely a performance limiting factor for the system and hence the SNR budget given for the phase noise is usually in a range of 30-40 dB at minimum and hence the overall system performance is limited by other factors. To give few numerical examples in the figure: at $f_c = 30$ GHz, the maximum bandwidth for 30 dB SNR_{pn} in the worse scenario (10 dB worse mobile terminal PLL) is 24 GHz, meaning that phase noise is not a limiting factor. However, for the same scenario the maximum bandwidth would be around only 950 MHz at $f_c = 150$ GHz and 240 MHz at f_c =300 GHz. These numbers are serious limiting factors. Same scenarios for 20 and 25 dB SNR_{pn}targets would be 9.5 and 2.4 GHz for f_c =150 GHz and be 2.4 GHz and 750 MHz at f_c =150 GHz. Moreover, even with symmetric link, assuming that both link ends have extremely good PLL performance, the phase noise is very drastically limiting the achievable signal bandwidth in high carrier frequencies.



(a) Single link end

(b) Whole link with equal PLL performances in Tx and Rx

(c) Whole link with 6 dB worse

(d) Whole link with 10 dB worse PLL performance

Figure 4-14: Phase noise-limited SNR. The different curves in the figure indicate the different phase noise limited SNRs.

Note that the analysis above assumed perfect cancellation of the lower frequency components of the phase noise. Hence, as some room must be left also for the compensation performance, the performance is decreasing further. Moreover, also the other RF nonidealities are contributing to the overall effective received SNR, the numbers given as an example gives rather pessimistic view for achieving very wide signal bandwidth in high frequencies at least with higher order modulations. However, the authors think that the problem can be solved in multiple domains and gives good research directions for further analysis and future work. First, it is evident that more SNR budget must be given for the phase noise at sub-THz region than in the lower frequencies. This is also natural as the target spectral efficiencies and lower order modulations are often versioned to be used that can survive with less signal to noise ratio. In waveform design, using analogue carrier-aggregation schemes with proper filtering could help to reduce the impact. There the trade-off is then to choose wide enough carriers to coop with lower offset frequency components of the phase noise but narrow enough carriers to tolerate the flat part of the phase noise. Such a channelized processing is also useful to generate and analyse signals by using multiple lower-bandwidth DACs and ADCs. Furthermore, different filtering schemes in OFDM-like waveforms may help to reduce the inter-symbol-interference caused by the wideband phase noise. Also, using analogue filtering to reduce phase noise in the PLL output could help to reduce the phase noise in some scenarios, but the filtering must be designed together with desired LO tuning range. Especially in those scenarios, understanding the noise budget in the multiplier chain and the overall noise level of the multipliers themselves would be critical.

4.2.2 Power amplifier modelling

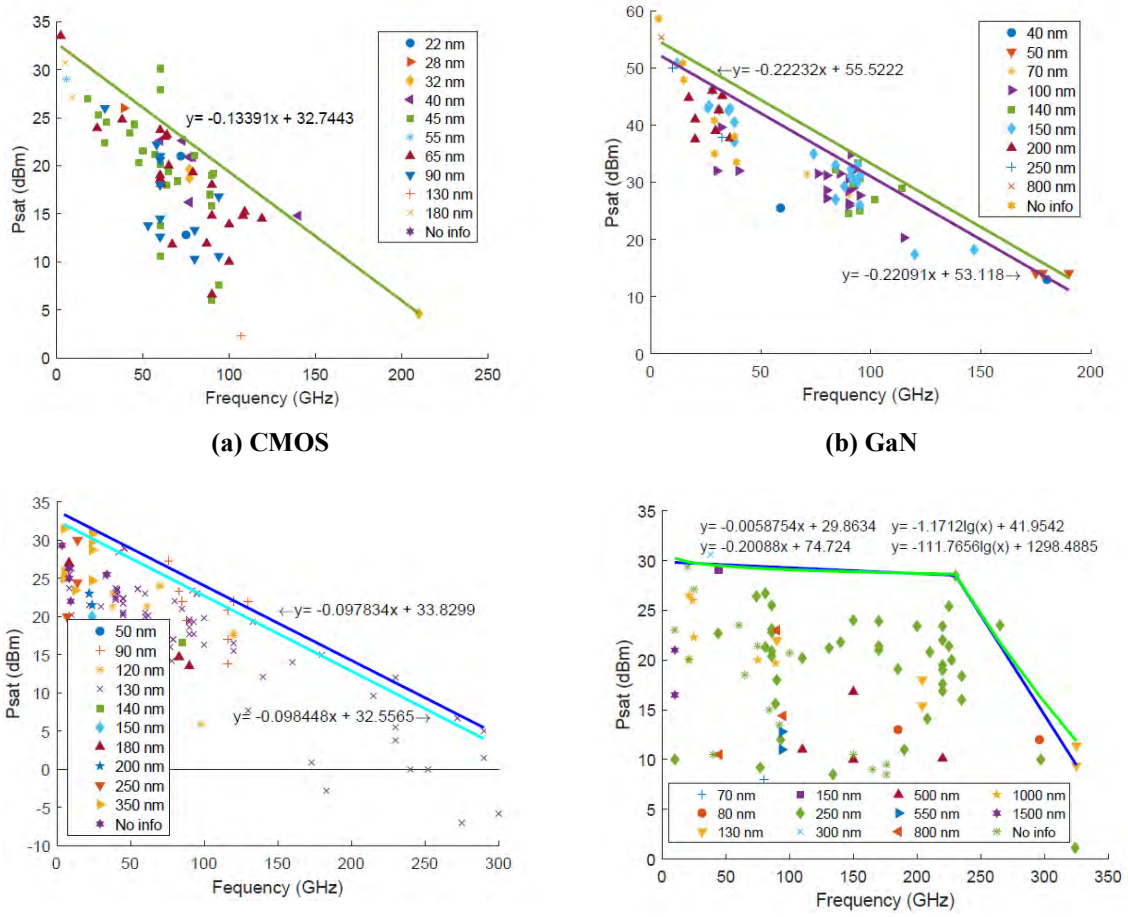
PA modelling is used to understand the nonlinear behaviour of the amplifier with different waveforms, linearize the PAs with different methods and approaches, as well as to analyse the achievable output power of the amplifiers to create realistic link budget. Adoption of upper mmWave frequencies challenge maximum power gain and oscillation frequencies, f_T and f_{max}, of the dominant circuit technologies are in a level of 200 – 500 GHz, depending on the foundry and technology selection [WWH+23]. Designing an amplifier operating 1/2 or even 1/3 of the f_{max} is extremely challenging and the circuit performance is inherently somewhat degraded compared to lower frequencies. At the same time, the used operating frequencies make it also extremely challenging to measure such circuits with simple measurement arrangements [LNT+21] and with good accuracy [LMN+22]. These challenges are also among the reasons why there is not publicly available PA model parameters for such frequency ranges.

This section is divided into three subsections: The first one analyses the realizable saturated power of different PA technologies. The aim is to give a scalable power model of different circuit technologies that can be then used when analysing link budget at different frequency bands. The second part focus on modelling the nonlinear behaviour of the PA by using different memoryless amplifier models. As an example, parametrization is given for a 300 GHz SiGe amplifier. In the third subsection, an approach for modelling the PA behaviour with memory effects is shown. The memory-model is based on physical causes of the memory effects such as gate–source capacitance, which are typically used in the device modelling of the transistor circuits.

4.2.2.1 Frequency scalable saturated power models with different semiconductor technologies

The 3GPP standardization has analysed the saturated RF power of PAs based on different semiconductor technologies for the proposed 5G frequency band from 7 to 24 GHz [38.820]. The basis for this in-depth analysis has been a comprehensive power amplifier survey performed by Georgia Tech [WWH+23]. The data survey consists of more than 3200 data points from publications collected over more than twenty years' time. It gives a good overview of the capabilities of different semiconductor technologies used for power amplifier purposes at different frequencies. The target of the analysis presented in this section is to present simple mathematical models of attainable maximum RF power based on the analysed data set to be used in 6G radio link-level simulations [RKL+20]. The analysed data set in the paper is based on analysis [WWH+23], and each imported data point has been verified from the original articles. All noticed errors have been corrected, and the duplicates have been removed. Only one highest-performing data point has been included from each design in the analysis. The analysed data set includes new added data points from the recent publications. The analysed data set has only one biasing or an operational condition to ensure the uniqueness of each data point in the analysis. The total number of amplifiers in the data set is 436 divided by technologies following: complementary metal-oxide semiconductor (CMOS) of 77, Gallium Nitride (GaN) of 73, SiGe of 125, Gallium Arsenide (GaAs) of 74 and Indium Phosphide (InP) of 87. From the data set four semiconductor technologies are presented in following analyses.

The saturated output power of CMOS amplifiers is shown in Figure 4-15 (a). For the CMOS A regression line approach was selected as a descriptive model since the best-performing amplifiers nicely formed a straight line when the frequency was on the linear, and the power on the dBm scale. The regression line is based on the fourteen highest performing amplifiers of the data set, and the coefficient of determination (R^2), which describes the goodness of fit, was 92.4% which indicates an excellent fit. Two outliers above the regression line can be seen from Figure 4-15, but those balance the regression line higher, and the effect of those is insignificant at 200 GHz. A 45 nm and 90 nm process nodes were dominant in the data set, while the highest performing PA used the 32 nm technology node.

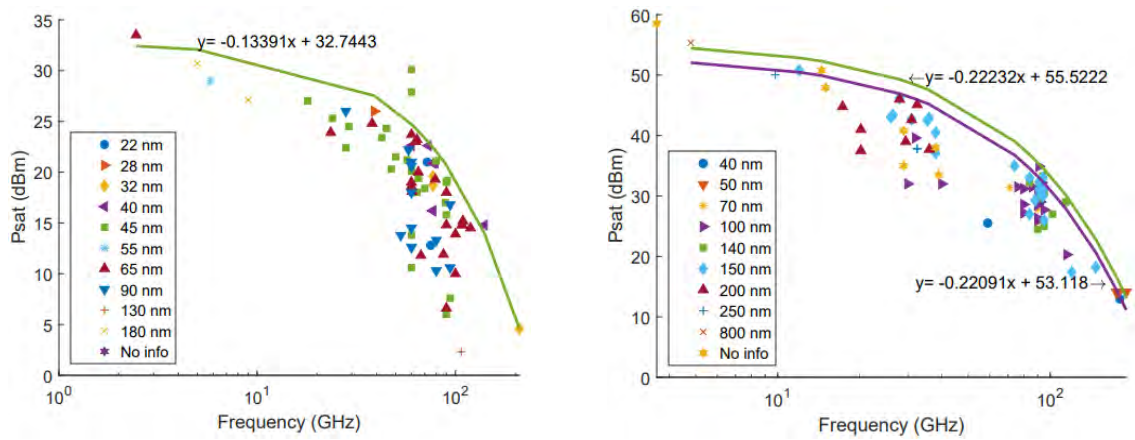


(a) CMOS **(b) GaN**

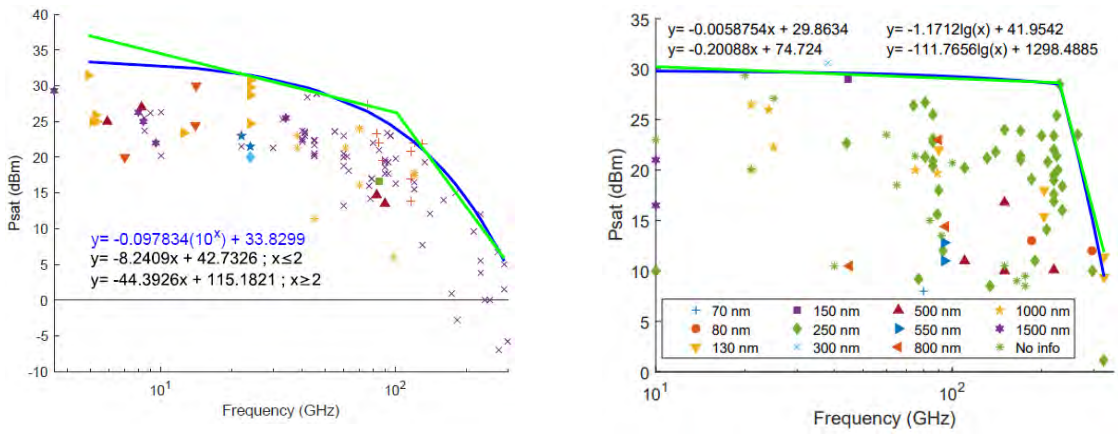
(c) SiGe **(d) InP**

Figure 4-15: Saturated output power models based on published PAs on linear frequency scale for different semiconductor technologies.

The InP-based power amplifier models are shown in Figure 4-15 (d) and Figure 4-16 (d). The dominant process node for the best-performing InP-based power amplifiers is 250 nm. The 130 nm node is used at the highest frequency PAs. The modelling of the saturated power of InP-based PAs requires two linear equations on both linear and logarithm scales. The threshold frequency of the modelling lines is 230 GHz since the slope changes from a flat (-1) to a very steep (-111) declining performance on the logarithm scale. The InP power modelling has similar curves regardless of the frequency scale (linear or logarithm).



(a) CMOS **(b) GaN**



(c) SiGe **(d) InP**

Figure 4-16: Saturated output power models based on published PAs on logarithm frequency scale for different semiconductor technologies.

4.2.2.2 Simplified memoryless amplifier models at 300 GHz

Different modelling approaches for different purposes and numerous different amplifier modelling approaches can be found from the literature all the way from various polynomial approaches [PM05] to neural networks [IBS+98]. These models achieve good performance, but on the other hand, they often require significant number of coefficients for which it is very challenging to find a generic parametrization that is scalable for different scenarios. In the system level, the simulations use often more simplified modelling approaches that aims for giving some insights on the amplitude-to-amplitude (AM-AM) and amplitude-to-phase (AM-PM) characteristics of a certain amplifier. Hence, different complexities of models are certainly required.

Four most used memoryless model approaches used in coarse system level evaluations are Rapp model [Rap91] with AM-PM extension [HH97], Ghorbani model [GS91], White model [WBJ03], and Saleh Models [Sal81, OML09]. All these models can be parametrized against different data, although some of them are specifically designed for a certain type of compression characteristics and PA topology. Parametrization of Rapp model has been also proposed for IEEE802.11 and 3GPP standards in [IEE09, IEE16, 3GR116, 3GR416] up to 60 GHz amplifiers. However, higher frequencies at upper mmWave region lacks such typical parameters This section present parameters for different well-known memoryless PA models against the amplifier design reported in [SRL+21] that operates at 300~GHz designed to operate above the $\frac{f_{max}}{2}$ of the used 130 nm SiGe technology. The AM-AM and AM-PM models can be generally presented as

$$y(x) = A(|x|)\exp\left(\frac{j\pi}{180}\varphi(|x|)\right)x, \tag{4-9}$$

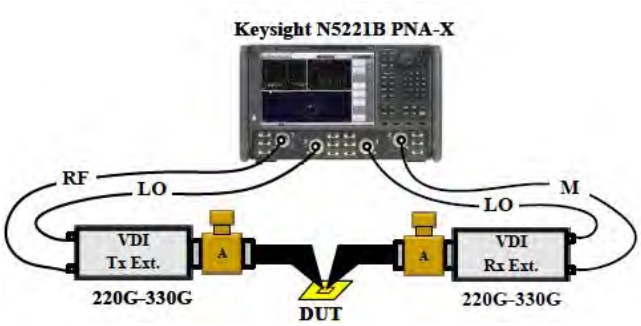
where x and y are the complex valued input and output amplitudes of the model, $A(|x|)$ represent the input-amplitude dependent gain of the amplifier and $\varphi(|x|)$ is the output-dependent phase shift of the amplifier (AM-PM) given in degrees. $A(|x|)$ and $\varphi(|x|)$ for modified Rapp [Rap91, HH97], Ghorbani [GS91], and Modified Saleh model [Sal81, OML09] are given in

Table 4-10. For White model [WBJ03], only AM-AM is given as it does not have AM-PM relation. For white model, one can use for example the AM-PM model of the modified Rapp model. Each of the parameter in the models have specific meaning available in the references and hence not discussed here in detail.

Table 4-10: Common memoryless nonlinear models of RF amplifiers.

Model	Mod. Rapp [HH97]	Ghorbani [GS91]	Mod. Saleh [OML09]	White [WBJ03]
$A(\|x\|)$	$\frac{g_r}{\left(1+\left(\frac{\|x\|}{x_{sat}}\right)^{2s}\right)^{\frac{1}{2s}}}$	$\frac{a_1\|x\|^{a_2}}{1+a_3\|x\|^{a_2}}+a_4\|x\|$	$\frac{a_s\|x\|}{1+b_s\|x\|^2}$	$a_w(1-\exp(-b_w\|x\|)$ $+c_w\|x\|\exp(-d_w\|x\|^2)$
$\varphi(\|x\|)$	$\frac{\alpha\|x\|^{q_1}}{1+\left(\frac{\|x\|}{\beta}\right)^{q_2}}$	$\frac{b_1\|x\|^{b_2}}{1+b_3\|x\|^{b_2}}+b_4\|x\|$	$\frac{c_s\|x\|}{1+d_s\|x\|^2}$	N/A

The measured and modelled 300 GHz amplifier used as a device under test (DUT) here is presented in detail in [SRL+21]. The amplifier is designed as an LNA, but the measured compression point of [SRL+21] is one of the highest reported saturated output power at that frequency implemented with the SiGe process. Thus, the amplifier is a PA, as well, when it operates near the saturation power levels. A new amplifier from the same manufacturing batch, as reported in [SRL+21] has been measured to derive the parameters presented in this section. There is some component-to-component variation that is typical for any high-frequency component. The measured saturated power is -3.6 dBm, which indicates a 5 dB range of the saturated power variation. The measured compression curve from the new chip is significantly smoother compared to results in [SRL+21]. The measured gain of 12.8 dB at 290 GHz is equal to the reported results.

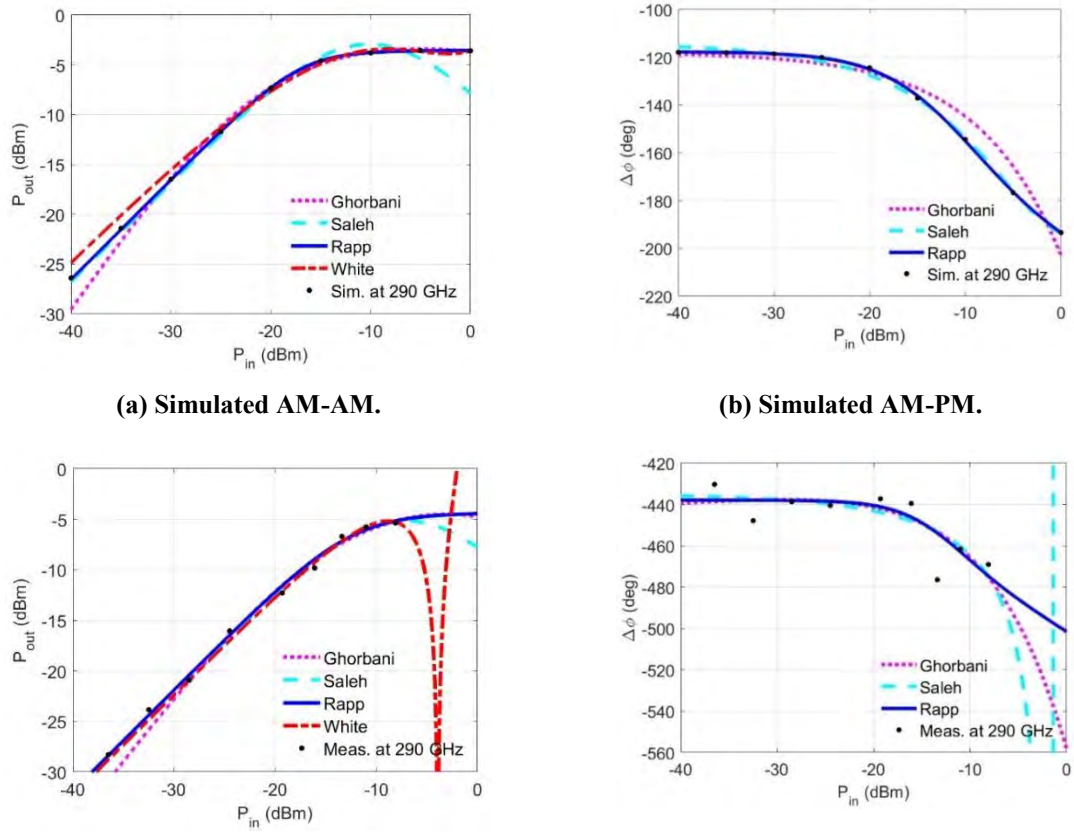


(a) Block diagram.

(b) Photograph of the laboratory arrangements.

Figure 4-17: Measurement setup.

A block diagram of the measurement setup for power-level-dependent scattering parameter (S-parameter) measurements is presented in Figure 5-17 (a), and a photograph of the system is shown in Figure 5-17 (b). The S-parameter measurement system includes two mechanically tuneable variable waveguide attenuators directly connected to the frequency extenders operating from 220 GHz to 330 GHz. The attenuator at the Tx side controls the input power to the DUT since the output power level of the Tx extender is constant. The attenuator at the Rx side prevents the overdrive of the Rx frequency extender enabling its linear operation. The RF measurement probes connect the DUT and the attenuators.



(a) Simulated AM-AM. **(b) Simulated AM-PM.**

(c) Measured AM-AM. **(d) Measured AM-PM.**

Figure 4-18: Performance of the 290 GHz SiGe amplifier, together with common memoryless amplifier models with fitted parametrization.

For each measured power level, an individual two-port calibration was performed. The simulated and measured power compression curve of the amplifier is shown in Figure 4-18 (a) and (c), and the phase plot over power levels is shown in Figure 4-18 (b) and (d) with black dots, respectively. The simulations are circuit-level simulations of the amplifier design. The measured AM-AM curve clearly follows a typical amplifier's compression curve. The variation can be seen from the AM-PM curve in the range of 20 degrees which corresponds to 55 um distance at 300 GHz. This variation is difficult to avoid when multiple mechanical probe adjustments are performed between calibrations and measurements at the 300 GHz range. Hence, the presented AM-PM here is having some uncertainty that cannot be avoided.

Based on the measured and simulated AM-AM and AM-PM at 290 GHz, all four typically used PA compression models were fitted to the datasets. The founded parameters are gathered to Table 4-11 and the corresponding curves are plotted to Figure 4-18. Note that the measured plots are purposely drawn over the fitting region to understand the behaviour outside the measurement range. In the models, amplitude is given normalized voltage (amplitude) against 1 Ω reference (can be used to calculate power directly as $|x|^2$), while AMPM parameters are given in degree scale. Figure 4-18 (c), Figure 4-18 (d) also contain the relative phase shift of the amplifier, which have been normalized away in the parametrization given in Table 4-11. Among all the models, Rapp model with gives the best compromise since has meaningful behaviour also outside the measurement range. This is important thing to keep in mind when performing system level simulations, as the model validity can be usually guaranteed only within the region to which it has been fitted. Both parametrizations, simulated and measured, for the Rapp model are given and may be used for system level simulations. Note that the parametrizations given here are not normalized and hence they also include the output power delivery capability of the amplifier under test.

Table 4-11: Parameter set for common memoryless amplifiers models fitted for circuit level simulations (sim) and measurements (meas) of 290 GHz SiGe amplifier.

Model (par).	AM-AM (par)	AM-AM (numeric)	AM-PM (par)	AM-PM (numeric)
Rapp (sim.)	{gr,xsat,s}	{4.708,0.663,1.603}	{α,q1,β,q2}	{−740.2,1.945,0.298,1.797}
Rapp (meas.)	{gr,xsat,s}	{2.55,0.605,1.296}	{α,q1,β,q2}	{−2.84 · 103,3.00,0.239,2.640}
Ghorbani (sim.)	{a1,a2,a3,a4}	{29.07,1.454,37.22,−0.1057}	{b1,b2,b3,b4}	{1.235,0.0018,−0.461,−111.856}
Ghorbani (meas.)	{a1,a2,a3,a4}	{7.628,1.296,8.742,−0.2037}	{b1,b2,b3,b4}	{0.4163,0.0013,−0.994,−157.4}
White (sim.)	{aw,bw,cw,dw}	{1.008,6.604,−0.733,0.760}	N/A	N/A
White (meas.)	{aw,bw,cw,dw}	{0.146,−3.456,20.159,1.156}	N/A	N/A
Saleh (sim.)	{as,bs}	{4.58,10.32}	{cs,ds}	{−132.85,0.673}
Seleh (meas.)	{as,bs}	{2.52,5.16}	{cs,ds}	{−79.85,−1.37}

4.2.2.3 Behavioural PA modelling with memory

PAs are essential building blocks in communication systems, especially in 5G and 6G communication systems where it is challenging to produce sufficient output power. With the extension of communication systems to higher frequencies, the signal bandwidth used in these systems will be comparable to the inherent bandwidth of PAs. This brings growing interest in PA baseband behavioural modelling that takes account both non-linearities and memory effects [SOG+08]. [OML09]. Baseband behavioural modelling is a result of observed PA characteristics using measured input and output baseband data. A complete measurement including all possible states of amplifiers by sweeping over the input power range and frequency range is a prerequisite for accurate PA behavioural modelling. However, it is difficult to measure the high-frequency broadband input and output and to work on PA behavioural modelling due to limited setups which by themselves may contribute with nonlinear effects.

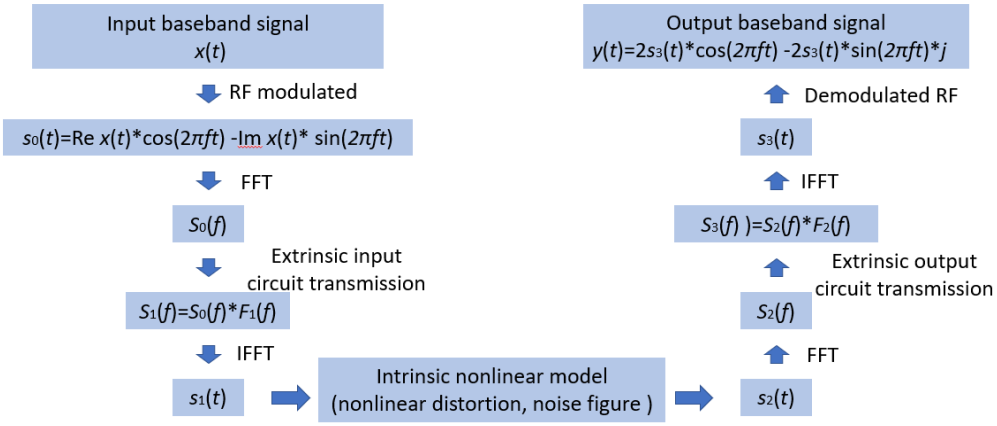


Figure 4-19: Structure for PA equivalent circuit simulation.

In this work a method to simulate the baseband output signal of a PA is presented using RF equivalent circuit simulation, in which memory effects and nonlinear effects are considered in the intrinsic circuit model simulation. Our method provides a new way for generating input-output baseband data that can be used for PA behavioural modelling. The validity of the methods is investigated in comparison with measured results of a GaN PA. Here, a method to simulate the baseband output signal of a PA is presented using RF equivalent circuit simulation, in which memory effects and nonlinear effects are considered in the intrinsic circuit model simulation. Our method provides a new way for generating input-output base-band data that can be used for PA behavioural modelling. The validity of the methods is investigated in comparison with measured results of a GaN PA.

The method depicted in Figure 4-19 is based on generating a RF signal using simulated base-band data (IQ samples). This RF signal is passed through an RF equivalent circuit simulation model of a PA. The resulting output signal of the circuit simulation is the down-converted to baseband again to produce the

base-band output signal. In this simulation approach, input and output mismatch network effects in the frequency domain are included.

The thermal noise, drain-current non-linearity and memory effects are introduced in the intrinsic circuit model, as listed below.

- Thermal noise: Random thermal agitation of charges inside a conductor produces thermal noise which is proportional to the temperature. This current voltage is given by Planck's blackbody radiation law [POZ11], $i_n = \sqrt{\frac{4hfB}{R(e^{\frac{hf}{kT}}-1)}}$, where h is Planck's constant, k is Boltzmann's constant, T is the temperature, B is the bandwidth of the system, f is the centre frequency of the bandwidth, and R is the device resistance.
- Drain current non-linearity: Non-linearities from drain current and gate–source capacitance can give rise to undesirable effects such as gain compression and the generation of spurious signals. Since the non-linearity effects of the drain current are much larger than those from the gate–source capacitance [VRD03]. Here, only the drain current non-linearity is considered in the PA simulation. The nonlinear current source controlled by the intrinsic gate voltage ($v_{gsi}\}$) can be modelled by the hyperbolic tangent (tanh) function [AZR92]: $i_d = i_0(1 + \tanh(p_1(v_{gsi} - v_{pk}) + p_2(v_{gsi} - v_{pk})^2 + p_3(v_{gsi} - v_{pk})^3))$, where $i_0$scales the current. v_{pk} and$p_{1,2,3}$ can be extracted by fitting v_{gsi}-$i_d$curves.
- Memory effects: The output of a PA with memory effects is not only determined from the instantaneous input, but also the input past and/or the system state [OML09]. Here only the memory effects from a linear gate–source capacitance are considered. In this case, the gate voltage is given by [VRD03], $v_{gsi}(t) = \frac{1}{C_{gs}}\int_{-\infty}^{t} i_{gs}(t')dt')$, where the voltage at time t is proportional to all prior gate current values, not just to the instantaneous value. Self-heating and trap-related mechanisms can also introduce memory effects. However, those are not considered in this work since some published measurement results show their minor dependence on memory effects [BG03, BP04].

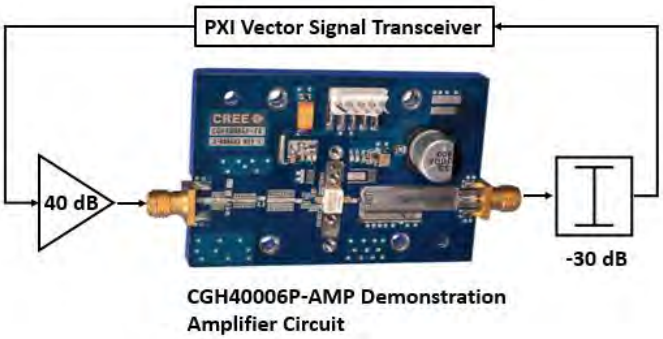


Figure 4-20: PA baseband measurement Setup [RFWL22].

To validate our method, a GaN PA (Cree CGH40006-TB) from Chalmers RF WebLab is used for demonstration. The input signal is a burst OFDM-like signal with different bandwidths (e.g., 10 MHz, 40 MHz and 160 MHz) and a carrier frequency of 2.14 GHz. Figure 4-20 shows the PA baseband input and output measurement setup which is based on PXI vector signal transceiver from National Instruments. Figure 4-21 shows the measured results and the circuit simulation results of the PA. The nonlinear fitting results without memory effects are also included for comparison. The results show that memory effects increase with the bandwidth, and the measured results are well predicted by the simulation results. The accuracy of the simulation results is evaluated by normalized mean square error (NMSE). Figure 4-22 shows NMSE between measured results and circuit simulation results as a function of bandwidth. The NMSE increases from -24 dB to -16 dB with bandwidths from 10 MHz to 80 MHz, and keeps stable at 80 MHz and 160 MHz.

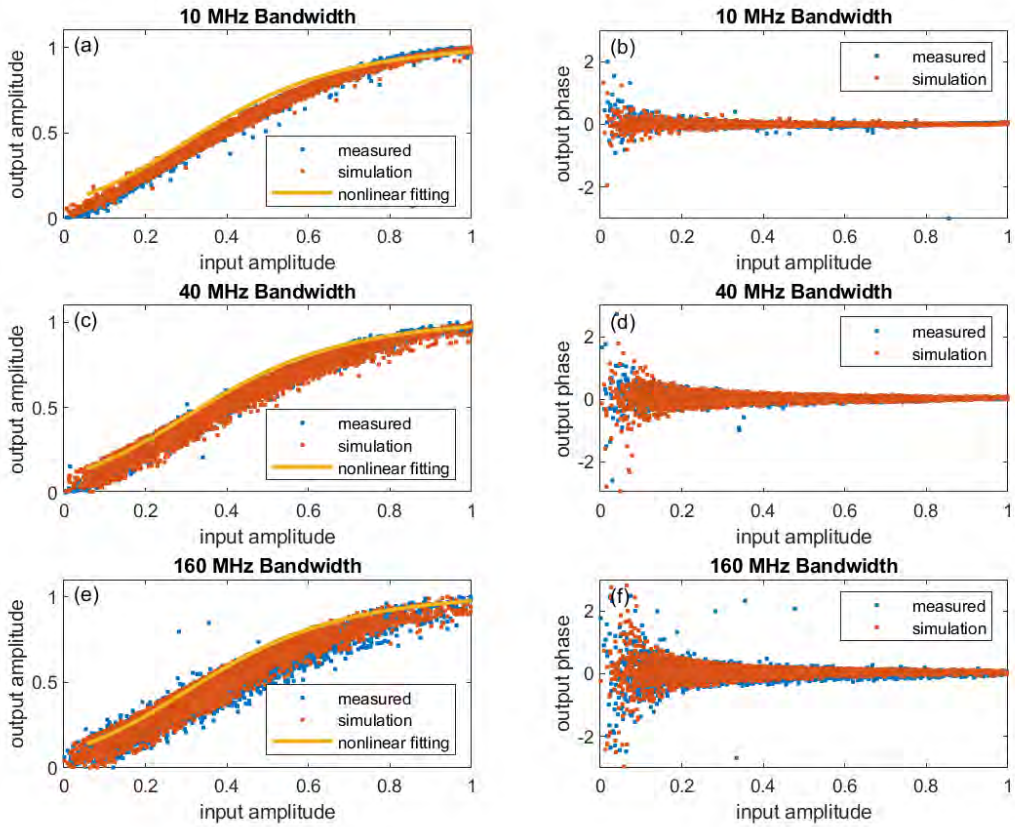


Figure 4-21: Measured and circuit simulation AM-AM (a, c and e) and AM-PM (b, d and f) results of the GaN with different bandwidths (10 MHz, 40 MHz and 160 MHz).

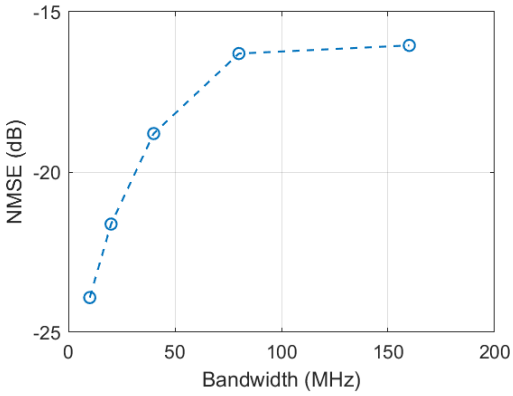


Figure 4-22: NMSE versus bandwidth.

Baseband memory polynomial model based on measured data in the discrete-time domain is widely used to describe nonlinear effects of PA with memory effects. The complex polynomial coefficients can be extracted by using the least-squares error method from input and output baseband results. Figure 4-23 shows the behavioural modelling results based on measured and simulation results with different bandwidths (10 MHz, 40 MHz and 160 MHz) of the GaN. The behavioural modelling results based on simulations follows the same trend as that based on measured results, while larger difference appears with 160 MHz bandwidth, which is due to lower accuracy of the simulation method, as described above.

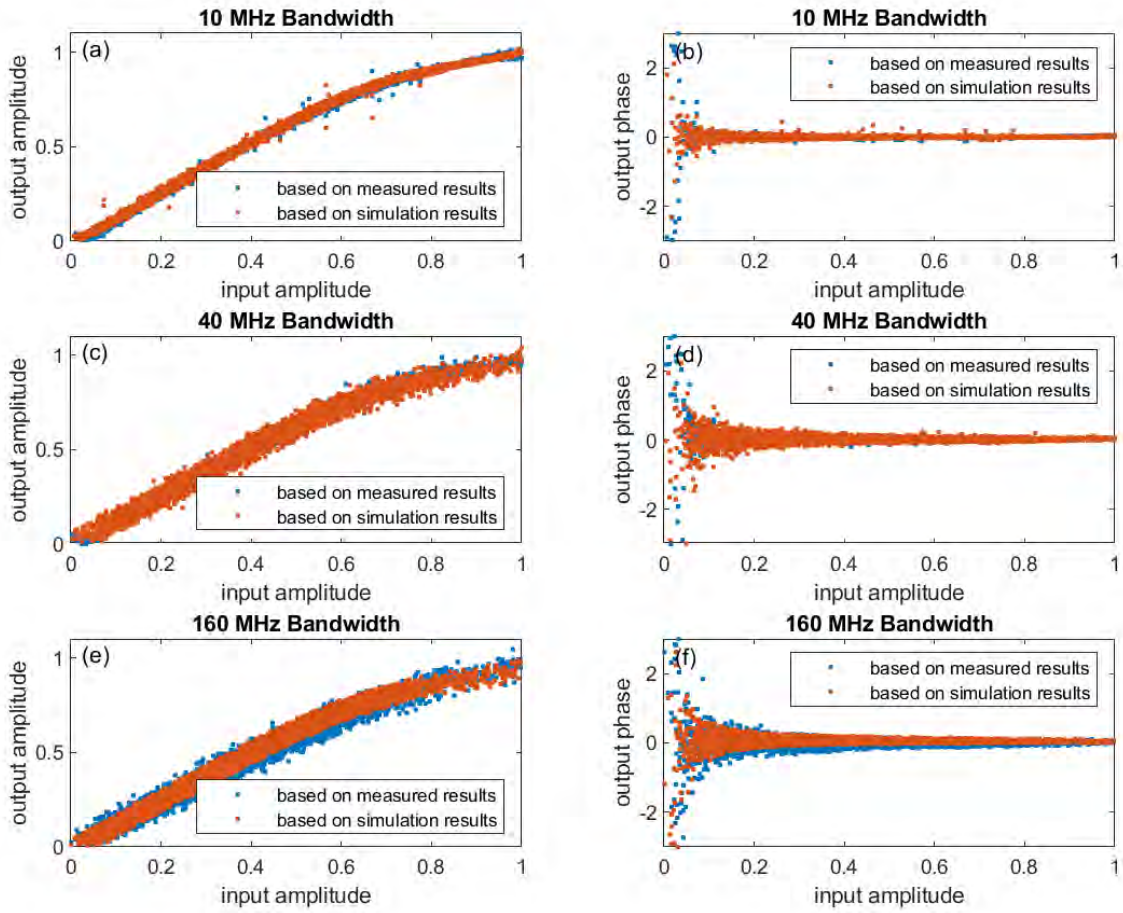


Figure 4-23: AM-AM (a, c, e) and AM-PM (b, d, f) behavioural modelling results based on measured and simulation results with different bandwidths (10 MHz, 40 MHz and 160 MHz) of the GaN.

In summary, RF equivalent circuit simulation method is demonstrated for generating PA baseband input output signal. The method including both memory effects and nonlinear effects is validated by GaN PA measured results. The NMSE between the measured results and circuit simulation results of the GaN PA is lower than -15 dB with bandwidth less than 160 MHz at 2.14 GHz. It is important to note that this method is not limited to GaN PA but also can be applied to any kinds of PAs. Similar PA modelling can be done for higher frequencies but needs appropriate measurement data for validation. More experimental data is needed to further validate the presented method for generating input-output baseband data.

4.3 D-MIMO radio architecture

4.3.1 Introduction

Any new 6G technology component needs to provide some significant functional benefits, such as delivering very high radio performance. While RAN performance to maximize traffic for a given site remains important, also other performance metrics, e.g., energy, transport efficiency, robustness, latency, mobility, security, deployment flexibility and footprint, become important for 6G network planning and deployment. Distributed MIMO (D-MIMO) is one of the techniques that can enhance the aforementioned metrics, by a high degree of macro diversity that results in a predictable service quality over the entire service area. The enabling techniques should support scalable D-MIMO systems in scenarios ranging from low bands to sub-THz bands taking deployment options into account. Tight interaction with legacy systems to secure efficient use of resources will also be needed. The integrated

solution together with reuse of already existing deployments will probably lead to that the interaction between legacy systems such that dense D-MIMO deployments can co-exist with legacy macro sites.

In dense urban area scenarios that requires high user experience data rates and high traffic area capacity, e.g., public squares, stadiums, airports, etc., D-MIMO networks can enhance performance with capacity-based features, e.g., higher spectral efficiency at frequency range 1 (FR1) bands; in conjunction with dense deployment on areas where capacity cannot be met by macro sites in an efficient way. It will provide robust deployments that carry decent traffic regardless of site location; enable efficient multi-user MIMO with coherent joint transmission with interference suppression capability. One can apply flexible total radiated power (TRP) selection/clustering based on the scenario requirements. Last but not least, outdoor D-MIMO installations need to have a low visual footprint due to invisible, and aesthetic impacts. With radio stripes, it is possible to hide the installation in existing construction elements in the environment.

For the indoor deployment scenarios Figure 4-24 e.g., factories, warehouses, offices focusing on dense machine-type communication (MTC) and extended reality (XR), while D-MIMO not only provides traffic area capacity, but also enhances reliability and resilience aspects. The line-of-sight probability for D-MIMO is very high which makes it suitable for deployment on very high frequency bands where radio propagation makes it challenging to provide a robust access link for mobile users. Phase coherent joint transmission is hard to realize at frequency range 2 (FR2) bands, but non-coherent joint transmission can help creating reliable access link due to macro diversity while it can still provide high rates thanks to high bandwidth despite its low spectral efficiency performance. Layer 1 (L1) mobility, e.g., combined cell type of radio units (RUs), which could be transparent to UE, can provide reliable and resilient mobility without frequent handovers. Besides, distributed environment inherently enhances network's sensing and positioning capabilities. In some scenarios, coverage from existing macro deployments may not exist. In these cases, D-MIMO networks can be used for data boosting but may require support for standalone D-MIMO operation.

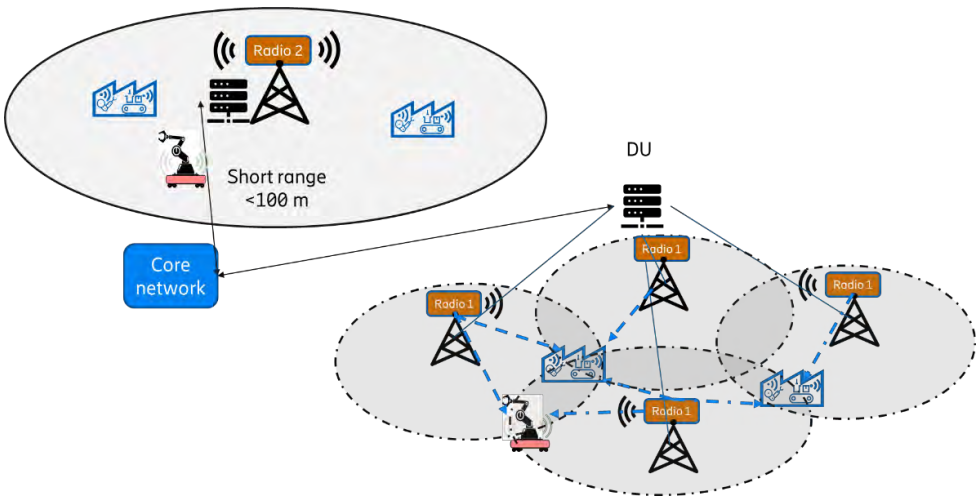


Figure 4-24 Macro base-station versus D-MIMO deployment.

The introduction of D-MIMO to a network means, that i) the network will have to support multi-connectivity of a user, ii) the network deployment will be densified and, iii) a much higher percentage of fronthaul links will be used in the network. These three aspects will not only affect deployment scenarios, signal processing, scheduling etc., but also affect user- and control-plane security. Multi-connectivity will need an arrangement for authentication and key agreement between the network and the user, which supports the connection of a user to multiple radio units. Deriving multiple keys and managing multiple secure connections may result in some extra overhead on the user side but in principle can be realized with today's techniques. D-MIMO schemes that require that the signal transmitted from multiple radio units combine coherently at the user will need a common encryption key to be used for the different signal paths. This applies to both user- and control plane traffic. Note

that in the control plane, not only the higher layers but also the MAC layer will have to be secured, so depending on the realization of the medium access control (MAC) layer in D-MIMO (how centralized it is) it might be necessary, that the network supports multiple security associations of a user, even if the higher layer control processing is centralized. To support the higher number of network nodes implied by D-MIMO, as well as multi-connectivity, it will be favourable to terminate user plane security more centrally in the network, for example where the application resides. This will reduce the number of exposed network points and will allow for using common user plane keys. The densification of the network using D-MIMO can be realized in different ways depending on how the functional split of radio unit (RU), distributed unit (DU), and centralized unit (CU) is implemented. In either case, it will be necessary that the nodes closer to the user, will be as cheap as possible (many nodes), and easy to deploy. Easy to deploy means the fronthaul nodes should support automatic certificate enrolment procedures such as specified for gNB for example. Based on this automatic enrolment, secure communication between the fronthaul nodes can be established.

If cheap and easy to deploy is the goal, this typically means that there must be a trade-off with other metrics. In this case, this means that cheap and easy to deploy nodes will not be as trustworthy as traditional eNBs or gNBs and not as resistant against physical attacks. The nodes will also be closer to the user, which makes them easier accessible for potential attackers. The network should not assume these entities are very trustworthy, and measures should be taken to detect compromise and malfunctioning and contain the impact of a compromised entity. In this context artificial intelligence (AI) may potentially be used in the future for monitoring and anomaly detection.

4.3.2 Architectures

D-MIMO systems can be deployed with different optimization goals in mind. At sub-6GHz, the reason for deploying advanced antenna solutions such as D-MIMO is mainly driven by the need for high-spectral efficiency. At very high frequencies, the need for distributed antenna systems is driven by the need to overcome the effects of a less reliable radio channel and to produce reliable communication links to the user equipment (UEs). Distributed systems allow for an architecture where a serving antenna is closer to a UE and thus can provide a more reliable link.

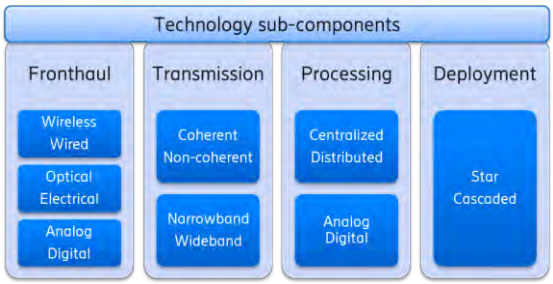


Figure 4-25: Illustration of architectural options for distributed MIMO.

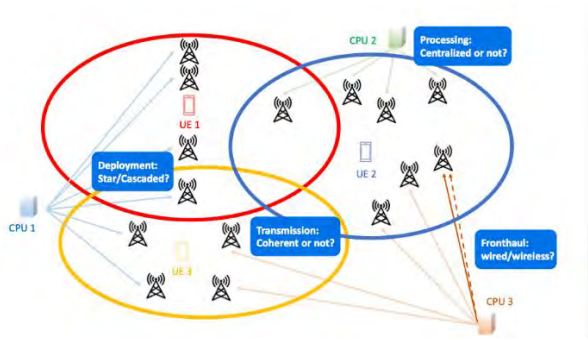


Figure 4-26: Illustration of distributed MIMO with key design options.

Higher frequencies and larger bandwidth in the access put higher demands on the front/backhaul for the distributed antennas as well as the processing of the high bandwidth signals. Processing and transmission can be categorized in the following ways:

1. Transport media: The backhaul/fronthaul transport media can either be wired or wireless.
2. Signalling: The data transmitted over the media can further be distinguished in being digitally encoded (e.g., common public radio interface (CPRI) like), or an analogue signal modulated onto a carrier.
3. Processing: Processing (such as for example beamforming) in different nodes can either be performed analogue or digitally. A further distinction can be made into centralized or distributed processing.

4. Transmission: If multiple serving antennas serve a UE jointly, this can either happen through coherent or non-coherent transmission by the serving antennas.

To enable scalable D-MIMO networks, there is a need to understand how much distribution of access points is needed and how to address practical issues associated with distributed access points operating at higher frequency bands. It is also necessary to understand the requirements on transport solutions and how these requirements can be met. The optimum solution would be phase-coherent joint transmission with centralized processing, but it is difficult to realize and might not meet the feasibility requirements. The other extreme would be phase non-coherent transmission with duplicating data in each RU. While this approach is technically feasible, it is inefficient. Furthermore, a combination of DMIMO and repeaters (more details in Section 4.3.3.2) will be an interesting topic to study. Not in all cases, one may require an RU, instead repeaters can extend the coverage in blind spots with relatively low cost.

The trade-off between localized (distributed) and centralized processing raises new challenges on the fronthaul (FH). The increased number of antennas for massive MIMO drives up the required FH load proportionally. There have been many FH challenges; some of them originated during 5G and some new types of FH challenges have been identified for new generations: further increased system bandwidth, e.g., at higher frequency bands, and multi-band support, more antennas, improved MU-MIMO performance with more layers, baseband resource pooling and FH statistical multiplexing, Inter-cell and interference coordination and joint processing, new D-MIMO schemes and deployments, e.g., star and serialized (cascaded) topologies. different lower layer split (LLS) options try to balance RU and FH complexity, cost and capacity for different use cases, while achieving high performance. The main differences can be summarized how to handle the following functionalities, e.g., where channel estimation is performed, where the beamforming weights (BFWs) are calculated, what types of beamforming types to support, which data, e.g., sounding reference signal (SRS), channel information, BFWs, etc., is transported over the FH interface, and their latency requirements.

Figure 4-27: D-MIMO with wireless fronthaul operating at high bands while access links at low bands.

There are two different ways to provide fronthaul to different RUs: wired through fibre, or via wireless links. An optimized wired fronthaul can be realized through so-called radio stripes, while wireless fronthaul/backhaul can be realized efficiently by integration with the access network. Cascading or serializing the fronthaul connections is also important for indoor installations, e.g., a factory environment, one fronthaul cable can connect several RUs or Antenna Processing Units (APUs) in some installations. Cable fronthauling is always preferable if it is easy, however wireless links are also an option for fronthauling, see Figure 4-25. At higher frequencies (e.g., sub-THz), the spectrum is large and there will not be much traffic in early phase of the utilization of these bands, even the UE is limited in how much bandwidth it can process, i.e., there will be available spectrum that can be utilized for the wireless fronthauling (WFH), Figure 4-27. Since RUs are generally not nomadic, robust fronthaul links will be easy to maintain. Random blockage due to moving objects can be handled by redundant links.

Moreover, in dense urban areas the backhaul connection for the DU can be provided wirelessly by an existing macro base station.

4.3.3 Analog approaches

4.3.3.1 Analogue over fibre based Fronthaul

Categorization: Media FB/BH Wired, Analogue, Central Processing: Digital.

In cases where large bandwidth signals are needed and/or low-complexity radio units are a must for cost-efficient deployments, current digital fronthaul interfaces face complications due to its limited spectral efficiency and flexibility. Analogue radio-over-fibre (ARoF) links are a prospective solution that can fulfil these needs, see Figure 4-28.

4.3.3.2 Network-controlled repeater and reconfigurable intelligent surface

Categorization: Media FB/BH Wireless, Transmission: Analogue, Node Processing: Analogue.

Network-controlled repeater (NCR) has been recently considered as one study-item in 3GPP Release 18 (finished in Aug. 2022), and the discussions are continuing in a work-item. In this section, the concept of NCRs is introduced as one possible low-complexity device to support for D-MIMO network densification and compare the performance of the NCRs with those achieved by reconfigurable intelligent surfaces (RIS). More technical details and results can be found in our submitted paper [GMM+22].

To improve the coverage and support the communication networks with massive numbers of UEs, different methods are proposed, among which network densification and mmWave communications are considered. Network densification refers to the deployment of multiple access points of different types in, e.g., metropolitan areas. Particularly, it is expected that in future small nodes, such as relays, integrated access and backhaul (IAB), repeaters, etc., may be deployed to assist the communications [REA+21]. In simple words, NCR is a normal repeater with beamforming capabilities and under the network control. Unlike reconfigurable intelligent surface (RIS) which controls the electromagnetic properties of the RF waves by performing an intelligent adaptation of the phase shift towards the desired direction, NCR is capable of advanced beamforming with power amplification.

As can be seen from Figure 4-29, the NCR-mobile termination (MT) is defined as a functional entity to communicate with a gNB via Control link (C-link) to enable the information exchanges, e.g., side-control information. The C-link is based on NR Uu interface. The NCR-forward (Fwd) is defined as a functional entity to perform the amplify-and-forwarding of uplink/downlink (UL/DL) RF signal between the gNB and the UE via backhaul link and access link. The behaviour of the NCR-Fwd will be controlled according to the side control information received by the NCR-MT from the controlling gNB. From a different perspective, RIS may be a possible technology for beyond 5G networks. It enables to control the electromagnetic properties of the RF waves by performing an intelligent adaptation of the phase shift towards the desired direction. In general, intelligent surfaces are electromagnetically active artificial structures with beamforming capabilities that can be used to reshape the propagation environment such as to improve capacity, coverage and energy efficiency. Similar to the NCR structure, RIS may also need a control link for configuration. RIS may be capable of signal reflection via adapting a phase matrix, while the NCR is capable of advanced beamforming with power amplification. As opposed to the NCR, RIS does not amplify the signal but also not the noise. Moreover, an RIS has only one beamforming matrix, whereas an NCR can do separate beamforming at receiver and transmitter side. This should be an advantage for the NCR when controlling the interference in, e.g., multi-user scenarios.

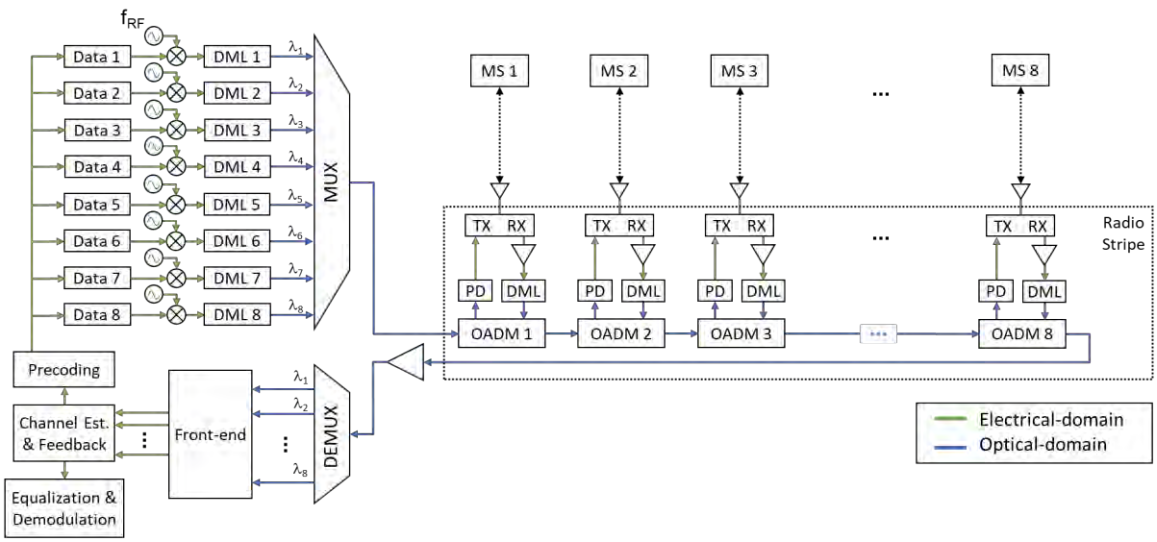


Figure 4-28: A wide bandwidth D-MIMO system can be implemented using ARoF and WDM.

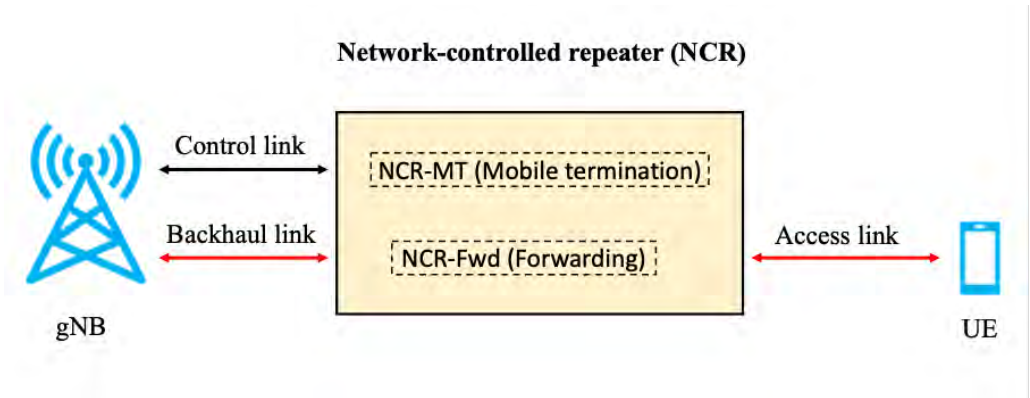


Figure 4-29: NCR structures [GMM+22, RP22673+22].

4.3.4 Digital approaches

Categorization: Media FH/BH Wired, Transmission: Coherent/Non-coherent, Processing: Centralized/Distributed.

Digital processing at high frequency bands, e.g., mmWave and sub-THz, will be power consuming and increases the chip cost and size of the processing unit. Small RUs are essential in terms of deployment aspects that is why it may be necessary to move the digital processing from RU to the distributed unit (DU), which means that FH will become analogue. On the contrary, in case of low frequency bands and limited bandwidth (e.g., 10-100 MHz), digital processing is feasible in the RU and can be carried out locally. Distributed (local) processing means that there will be less load on the FH since baseband data will be transferred, which can be handled without high-capacity FH links. However, in case of centralized processing supported by advance coordination to obtain better performance, the RUs will be less complex and the information should be conveyed to the DU. A smart way is required to find the balance in terms of complexity versus robustness and performance.

In D-MIMO systems, the dynamism could come from, e.g., mobility of users and changing of propagation environments. It may affect the system performance drastically, especially at high frequencies. Due to increasing distance and the influence of blockage, the channel of mmWave and sub-THz channels in a D-MIMO system decays rapidly. Consequently, there are several links between users and RUs that are too weak, so they can potentially be disregarded, resulting in a sparse channel for each user. Consider a dynamic D-MIMO system, in which the connections between users and RUs could vary as a user changes location or moves out of or into the served area. In this way, the dynamic D-MIMO problem can be represented as the followed figure. In order to simplify the dynamic topology

structure model, two types of links are considered in this work: strong and weak links. It should be noted that the strong or weak link can be defined by a threshold of channel strength, e.g., the distance, or path loss between UE and RU. Strong links must be estimated, whereas weak links can be ignored, but they have an impact on the pilot contamination. Figure 4-30 (a) indicates the *t*-th snapshot, while Figure 4-30 (b) indicates the (*t*+1)-th after structural modification. In comparison to the previous snapshot, in these figures the red and black lines represent new and lost connections, respectively.

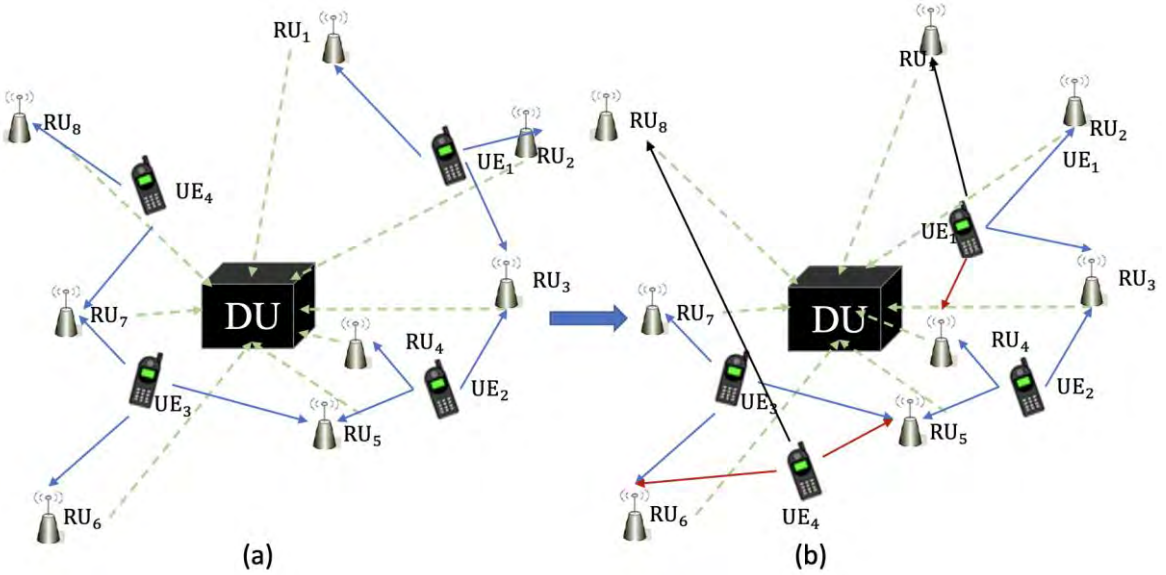


Figure 4-30: Dynamic D-MIMO model: connection modification.

4.3.5 Integrated access and backhaul

Categorization: Media: FB/BH Wireless, Transmission: Wideband, Processing: Centralized/Distributed.

The overall IAB architecture is based on the Centralized Unit (CU)/ DU split in the gNB. Thus, IAB donor consists of the CU and DU functionalities and connects to the core network using non-IAB type traditional backhauling, e.g., fibre, backhaul. Two IAB architectures are considered, namely, i) A hierarchical or, a cyclical structured IAB architecture where it is well-defined if a certain node is "above" or "below" a certain other node. Also, in such network the information flows in well-defined downlink and uplink directions. The child IAB nodes include two modules, namely, the DU and mobile termination (MT). Here, the DU serves the UEs in the access links as well as, potentially other downstream child IAB nodes in multi-hop wirelessly backhauled IAB networks. On the other hand, the MT is the module which connects the child IAB node with the DU of the parent/upstream node. ii) Mesh-like IAB architecture with no such well-defined hierarchy. Mesh architectures may provide efficient link scheduling capabilities and traffic adaptability, especially, in the presence of burst traffic in the network. Also, when moving to higher carrier frequencies, links tend to be less reliable and have a shorter. Mesh architectures in these cases can help to mitigate some these effects, although at the cost of increased cost and complexity. Here, the MT and DU split is replaced by the peer unit (PU) that can switch its mode into Master and Slave by the IAB donor determined by a scheduling scheme. The child IAB nodes in Master mode can control the backhaul links associated with the Slave mode in addition to the access links.

Figure 4-31 (a) indicates a planned hierarchical architecture where the child IAB nodes include two modules, namely, DU and MT while Figure 4-31 (b) indicates a mesh-like architecture which make the adjacent IAB node relationships dynamically configurable via the PU.

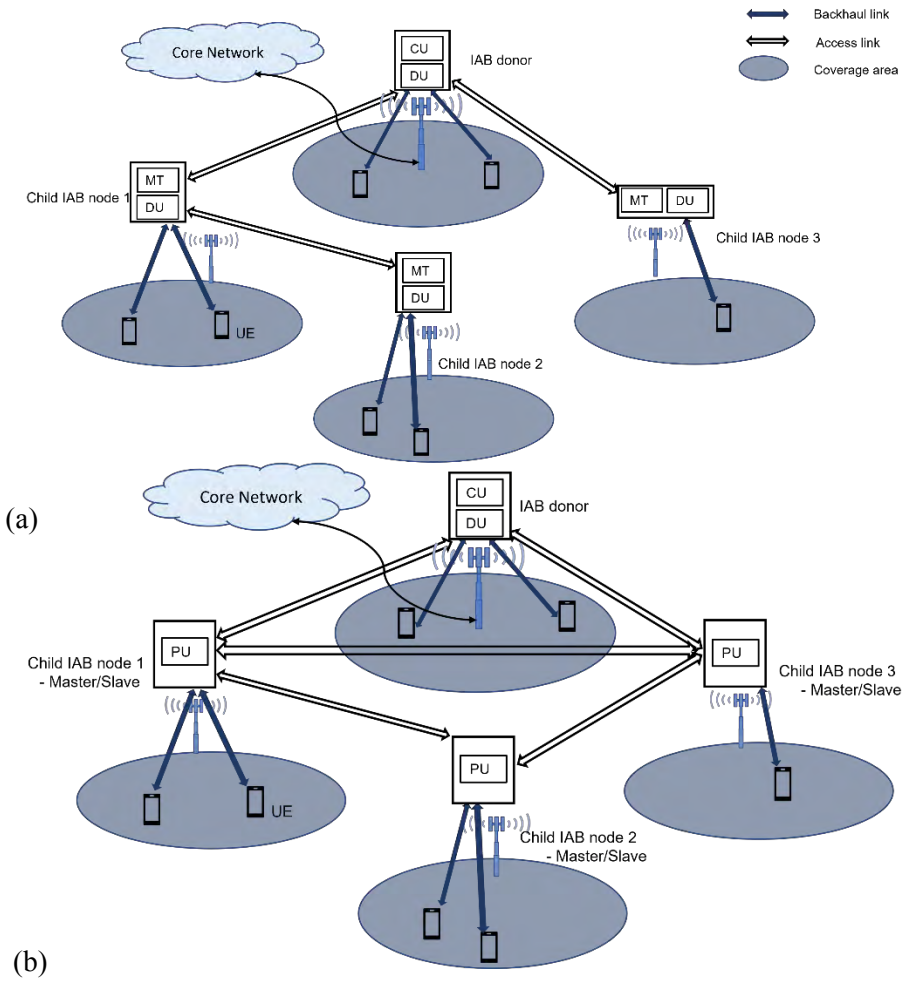


Figure 4-31: An illustration of IAB network (a): A planned IAB network architecture based on a hierarchy, (b) An IAB mesh network architecture.

5 Signal processing techniques

This section provides a detailed look into a selection of signal processing topics related to the design of high-throughput 6G systems. In the first part, an overview is given of selected waveform design topics and related transceiver design guidelines. In the second part, guidelines are given on how the beam management techniques need to be adapted to the large number of beams needed in sub-THz systems. The third part offers a look into a selection of topics from the domain of distributed MIMO (D-MIMO) and integrated access and backhaul.

5.1 HW-aware waveform and baseband transceiver design

Design of waveforms under the impact of hardware constraints as well as system design constraints (latency) is analysed in this section. Specifically, three waveforms are considered for a closer investigation: DFTS-OFDM (DFT-spread OFDM), SC-FDE (single-carrier with frequency domain equalization) and ZXM (zero-crossing modulation). Some qualitative aspects of these waveforms are given in Table 5-1, which forms part of a more detailed analysis in [HEX21-D21, Section 3.4.4].

Table 5-1: Qualitative assessment of selected waveforms.

	ZXM	DFTS-OFDM	SC-FDE
Power efficiency	Medium	Medium	Medium
Spectral efficiency	Medium to low	High	High
Robustness to HW impairments	Open problem	Medium to high	Medium to high
Implementation complexity	Low	Medium	Medium
Compatibility with MIMO	High	Medium	Medium
Flexibility of multiuser allocation in frequency	Low	High	Medium to low
Robustness to time-selective channels	Open problem	Medium	Medium
Robustness to frequency selective channels	Open problem	High	High

5.1.1 User-plane latency analysis

One of Hexa-X project objectives/visions is achieving a radio access network (RAN) latency of 0.1 ms [HEXA-XD2.1, Table 2.1]. In this contribution, it is defined what is meant by RAN latency and the feasibility of achieving the proposed target is analysed.

In this document the downlink (DL) latency is analysed. The related uplink (UL) latency is in addition to the aspects of the DL latency also more dependent on how the system is scheduling UL transmissions. Therefore, the DL latency can be seen as the lower bound of the UL latency.

Definition
Under RAN latency, **user plane latency** is assumed**.** As a working definition of RAN latency, the 3GPP definition of user plane latency may be adopted:

"The time it takes to successfully deliver an application layer packet/message from the radio protocol layer 2/3 SDU ingress point to the radio protocol layer 2/3 SDU egress point via the radio interface in both uplink and downlink directions, where neither device nor Base Station reception is restricted by DRX."

NR frame structure extrapolation

Establishing the time structure of layer-1 transmissions is essential for solving the problem of latency evaluation. The analysis assumes an NR-like time structure based **on OFDM symbols with cyclic prefix**, arranged in **slots**. Time/frequency resources in a slot are assigned to both data and reference signalling; hence, a slot is the basic unit of processing time. Slot-based time structure applies to OFDM and OFDM-like waveforms, such as DFTS-OFDM; however, it can be assumed that other waveforms, such as pure single carrier, can be represented by a slot-like structure, where data and reference signals needed for correct reception of data are e.g. time-multiplexed.

As another set of working assumptions include

- Extrapolation of OFDM symbol time durations in NR for larger than NR sub-carrier spacing (SCS) configurations. Higher SCS are tightly connected to large data bandwidths and high sampling rates; therefore, the SCS can be seen as a measure of bandwidth and sampling rate and can in this context be **assumed to represent general waveforms, not only OFDM-like** waveforms.
- Adoption of the number of (OFDM) symbols in a slot as a parameter for the parametric latency calculation. In the results, **1-symbol** and **2-symbol slots** are assumed. Note that 2-symbol mini slots are also present in NR.

Numerologies/subcarrier spacings, corresponding OFDM symbol durations and corresponding maximum supported BW (for standard FFT size of 4096 and not assuming frequency guard bands) are shown in Table 5-2. As a clarification: a system under analysis can be based on OFDM-like waveforms with **symbols** containing (4096 + 288) samples, with 1 or 2 such symbols forming a slot, or by considering a system that usees the same sampling rate but a different waveform, where samples of this waveforms form **subgroups** of (4096 + 288) samples each, and **groups** formed by 1 or 2 subgroups. The latency analysis for a given numerology applies to both systems equally, given that they use the same sampling rate, which is represented by the numerology.

Table 5-2: Extrapolation of NR numerology for FFT size 4096.

sub-carrier spacing	OFDM symbol duration + CP, in µs	Maximum supported BW = sampling rate (oversampling ratio = 1)
µ = 0 (SCS = 15 kHz)	71.4	61.44 MHz
µ = 1 (SCS = 30 kHz)	35.7	122.9 MHz
µ = 2 (SCS = 60 kHz)	17.85	245.8 MHz
µ = 3 (SCS = 120 kHz)	8.925	491.5 MHz
µ = 4 (SCS = 240 kHz)	4.4625	983 MHz
µ = 5 (SCS = 480 kHz)	2.2313	1.96 GHz
µ = 6 (SCS = 960 kHz)	1.1156	3.93 GHz
µ = 7 (SCS = 1920 kHz)	0.5578	7.86 GHz
µ = 8 (SCS = 3840 kHz)	0.2789	15.73 GHz

Another option is to further increase the SCS by limiting the FFT size to 1024 or 2048 sub-carriers. The feasibility of further scaling is depending on the delay spread of the channel. Thus, current, and future investigation of the channel characteristics with and without beamforming will show whether this is feasible for an OFDM like system. To compensate for the effect of delay spread, it is also possible to increase the cyclic prefix (CP) overhead. Current new radio (NR) systems (with normal CP) keep a constant CP overhead of about 7%. The result of this extrapolation is shown in Table 5-3.

Table 5-3: Extrapolation of NR SCS for FFT size of 1024 and 2048.

sub-carrier spacing in kHz	OFDM symbol duration + CP, in µs	FFT size 1024 BW excluding guard band (~60 % SCs used)	FFT size 2048 BW excluding guard band (~60 % SCs used)
15	71.4	9.21 MHz	18.4 MHz
30	35.7	18.4 MHz	36.8 MHz
60	17.85	36.8 MHz	73.7 MHz
120	8.925	73.7 MHz	147 MHz
240	4.4625	147 MHz	295 MHz
480	2.2313	295 MHz	589 MHz
960	1.1156	589 MHz	1.18 GHz
1920	0.5578	1.18 GHz	2.36 GHz
3840	0.2789	2.36 GHz	4.72 GHz
7680	0.13945	4.72 GHz	9.43 GHz

Following the choice of RAN latency definition, the evaluation procedure for determining user plane latency proposed by 3GPP for the NR standard [37.910, Section 5.7] may also be adopted. Under this definition, the possibility of employing retransmissions to achieve a **satisfactory performance level** is accepted. Evaluation of user plane latency between user equipment (UE) and base station (BS) is illustrated in Figure 5-1, taken from [37.910, Section 5.7.1.1].

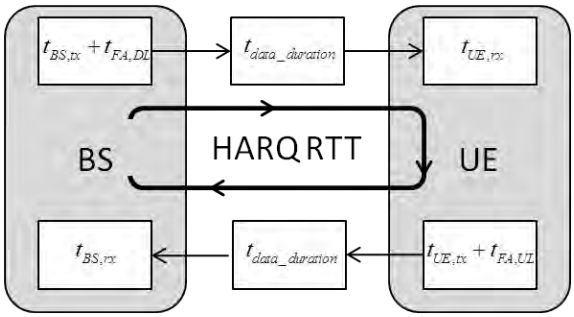


Figure 5-1: User plane procedure for latency evaluation.

Definition of different terms contributing to latency is provided in Table 5-4.

Table 5-4: Terms for user plane latency evaluation.

Value	Definition	Description
$t_{BS,tx}$	BS processing time/delay	Time interval between the arrival of data at layer 2/3 ingress point and generation of physical layer packet
$t_{FA,DL}$	DL frame alignment time	Expected waiting time for the start of next slot
$t_{data\_duration}$	Transmission time interval (TTI) for DL data packet transmission	
$t_{UE,rx}$	UE processing time/delay	interval between the time when DL transmission is received, and data is decoded
$t_{UE,tx}$	UE processing delay	interval between the time the data is decoded and the time when ACK/NACK packet is generated
$t_{FA,UL}$	UL frame alignment time	Expected waiting time for the start of next slot
$t_{BS,rx}$	BS processing delay/time	interval between the time when ACK/NACK is received and the point when ACK/NACK is decoded

DL transmission time is

$$T_1 = t_{BS,tx} + t_{FA,DL} + t_{data\_duration} + t_{UE,rx}.$$

UL transmission time (for sending acknowledgment (ACK)/NACK packets in a retransmission) is

$$T_2 = t_{UE,tx} + t_{FA,UL} + t_{data\_duration} + t_{BS,rx}.$$

Total transmission time with n_{RT} retransmissions is:

$$T_{total} = T_1 + n_{RT}(T_1 + T_2).$$

UE and BS processing time dependent on slot duration

In the calculation above, special care is devoted to processing delays $t_{BS,tx}$, $t_{BS,rx}$, $t_{UE,tx}$ and $t_{UE,rx}$. In the first part of the analysis, it is assumed that UE and BS processing time depends on the slot duration and distinguish 2 separate cases:

1. $t_{BS,tx} = t_{BS,rx} = t_{UE,tx} = t_{UE,rx} = 1$ slot. Assuming that one layer-2 packet corresponds perfectly to 1 slot, this is e.g. the minimum time needed for encoding/decoding the error correction code and possibly performing a cyclic redundancy check (CRC) calculation/check. Also, this is the minimum time needed for performing channel estimation and equalization and applying the channel estimates. This case can be labelled as "high processing capability" case.
2. $t_{BS,tx} = t_{BS,rx} = t_{UE,tx} = t_{UE,rx} = 2$ slots. Here a layer-2 packet corresponds to up to 2 slots, and/or processing times are non-negligible. This case can be associated with the label "low processing capability".

Values of terms from Table 5-4 used in the analysis are elaborated in Table 5-5.

Table 5-5: Assumed values of parameters for latency calculation, processing time tied to slot duration.

Value	Definition	Description
$t_{BS,tx}$	BS processing time/delay	a) 1 slot b) 2 slots
$t_{FA,DL}$	DL frame alignment time	slot duration/2 (assuming uniform distribution from 0 to slot duration)
$t_{data\_duration}$	TTI for DL data packet transmission	1 slot = a) 1 OFDM symbol, b) 2 OFDM symbols
$t_{UE,rx}$	UE processing time/delay	a) 1 slot b) 2 slots
$t_{UE,tx}$	UE processing delay	a) 1 slot b) 2 slots
$t_{FA,UL}$	UL frame alignment time	slot duration/2 (assuming uniform distribution from 0 to slot duration)
$t_{BS,rx}$	BS processing delay/time	a) 1 slot b) 2 slots

Results of the parametric analysis are shown in Figure 5-2 to Figure 5-5. It can be seen that a higher number of retransmissions requires a shorter slot duration (in turn resulting in the use of a higher subcarrier spacing/bandwidth) to meet the latency requirement of 100 μs. This analysis, however, assumes that processing time scales with slot size, which may not be the case in practice. To this end, a parametric analysis of latency with processing time independent of slot duration is needed, and this analysis is presented in the following.

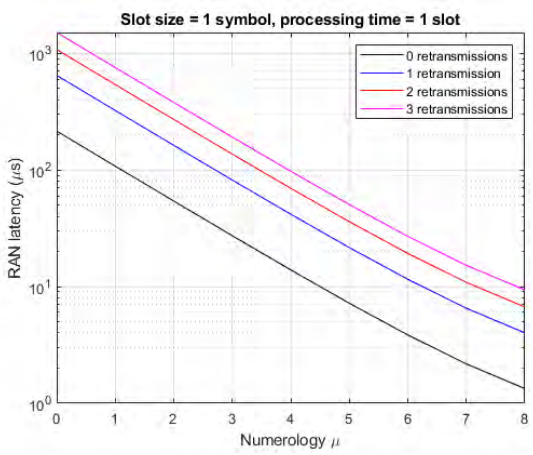


Figure 5-2: RAN latency for 0-3 retransmissions, slot size of 1 OFDM symbol and UL/DL BS/UE. processing times of 1 slot (high processing capability).

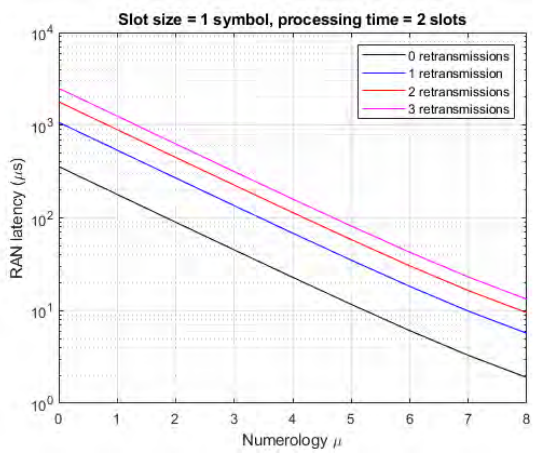


Figure 5-3: RAN latency for 0-3 retransmissions, slot size of 1 OFDM symbol and UL/DL BS/UE. processing times of 2 slots (low processing capability).

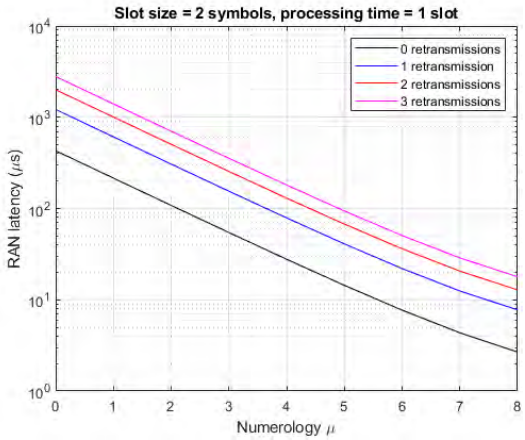


Figure 5-4: RAN latency for 0-3 retransmissions, slot size of 2 OFDM symbol and UL/DL BS/UE processing times of 1 slot (high processing capability).

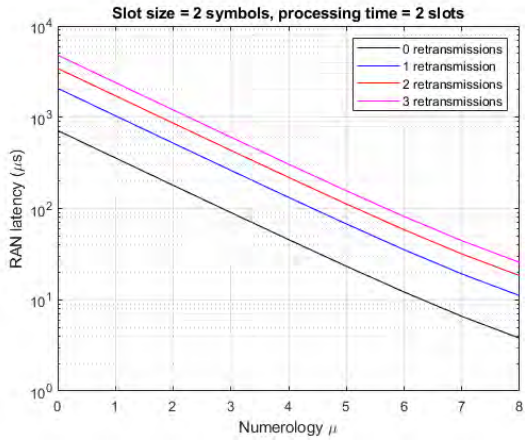


Figure 5-5: RAN latency for 0-3 retransmissions, slot size of 2 OFDM symbol and UL/DL BS/UE processing times of 1 slot (high processing capability).

UE and BS processing time independent of slot duration

The first analysis is keeping the SCS constant and the values for $t_{BS,tx}$, $t_{BS,rx}$, $t_{UE,tx}$, and $t_{UE,rx}$ are parameterized and have the following relationship

$$t_{BS,tx} = t_{BS,rx}, \quad and \quad t_{UE,tx} = t_{UE,rx}.$$

Also, in the first part of the analysis, processing times shrink together with slot size, leading to a processing time of about 1 µs for SCS = 3840 kHz, 2-symbol slots and processing time of 2 slots. It is unclear if processing times of this order of magnitude are feasible and if yes, under what parallelization and power consumption cost. In the second part of the analysis, the processing time is decoupled from slot duration and parameterized as a fraction of state-of-the art processing time. Namely, the lowest values defined for the evaluation in [37.910] serve as the upper bound for these values assuming 60 kHz SCS. This means that the maximum value of $t_{UE,tx}$ is defined as 98.2 µs and $t_{BS,tx}$ as 80 µs [38.214, Table 6.4.-2 and Table 5.3.-2]. As it is expected that these processing times evolve with the same pace, a scaling parameter α is introduced for both of these parameters to scale them relative to each other. This means for the results in this section $t_{UE,tx} = \alpha\, 98.2\ \mu s$ and $t_{BS,tx} = \alpha\, 80\ \mu s$.

Results of the parametric analysis are shown in Figure 5-6 to Figure 5-8**.**

Table 5-6: Assumed values of parameters for latency calculation, processing time independent of slot duration.

Value	Definition	Description
$t_{BS,tx}$	BS processing time/delay	parameterized
$t_{FA,DL}$	DL frame alignment time	slot duration/2 (assuming uniform distribution from 0 to slot duration)
$t_{data\_duration}$	TTI for DL data packet transmission	1 slot = 2 OFDM symbol, for uplink 1 OFDM symbol
$t_{UE,rx}$	UE processing time/delay	parameterized
$t_{UE,tx}$	UE processing delay	parameterized
$t_{FA,UL}$	UL frame alignment time	slot duration/2 (assuming uniform distribution from 0 to slot duration)
$t_{BS,rx}$	BS processing delay/time	parameterized

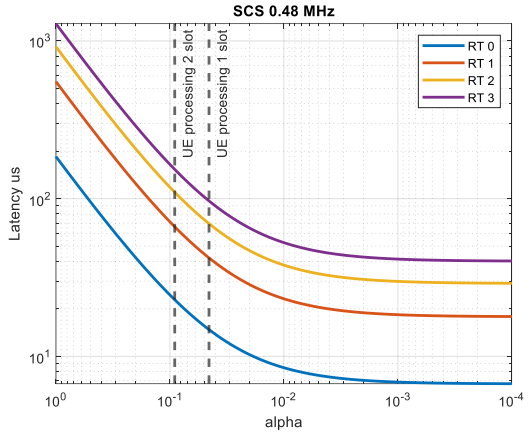


Figure 5-6: Processing time scaling for SCS of 480 kHz.

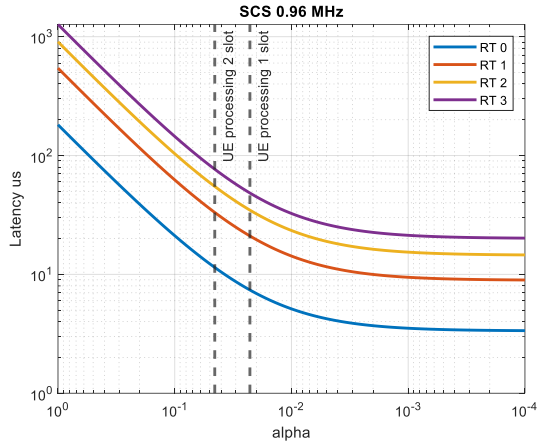


Figure 5-7: Processing time scaling for SCS of 960 kHz.

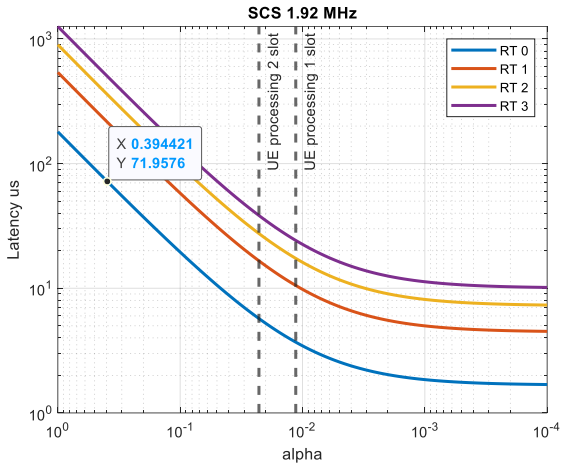


Figure 5-8: Processing time scaling for SCS of 1.92 MHz.

Discussion

From the results of the parametric analysis, it is evident that the RAN user plane latency of 0.1 µs is generally achievable at numerologies $\mu = 5$ and above (corresponding to bandwidths of 2 GHz and above), if a) slot size is kept low (1 or 2 OFDM symbols) and b) processing times are on the order of 1-2 slots. In this case, retransmissions are possible. Latency, naturally, degrades with a larger number of retransmissions and longer processing times (for slot = 2 symbols and processing time of 2 slots a signal BW of 4 GHz is needed for meeting the latency requirement with 3 retransmissions).

Regarding the second part of the study, the results in in Figure 5-6 to Figure 5-8 show that for the used slot structure the latency target of 100 µs can be met if the processing times at UE and BS scale down by about an order of magnitude compared to state-of-the-art for SCS = 60 kHz. In nominal terms, this means 10 – 50 μs of processing time per slot at both UE and BS side. Also, it is worth noting that the latency target in this case can be met with significantly higher processing time than 2 slots. This is dependent on the SCS but assuming that a maximum of 1 retransmission is used to achieve the required reliability a processing time of about 15 µs is sufficient.

For finding the ultimate answer on whether meeting the proposed RAN latency requirement is feasible, it is necessary to find out reasonable UE and BS processing times, considering the projected computational capability scaling, I/O speeds, memory access speeds and power consumption constraints.

5.1.2 SC-FDE and DFTS-OFDM: parameterization and performance

5.1.2.1 Impact of HW impairments on the performance of SC-FDE and DFTS-OFDM

5.1.2.1.1 Impact of phase noise

This part focuses on investigating the impact of the non-linear distortion of the power amplifier (PA) at the transmitter with the phase noise (PN) at the receiver in a DL scenario. The PN model is the one that is described in Section 4.2.1.1 under model 2, being scaled to 140 GHz carrier frequency. The PA model is the same as the one used in [HEX21-D22, Section 3.6.6.1]. This investigation is comparing DFT Spread OFDM (DFTS-OFDM) with Single-Carrier Frequency Domain Equalization (SC-FDE). The simulation parameters are described in Table 5-7.

Table 5-7: System parameters for DFTS-OFDM and SC-FDE comparison study.

RF impairment	PA nonlinearity
PA model	memoryless GaN
PN model	Model 2 described in Section 4.2.1.1 scaled to 140 GHz
Carrier frequency	140 GHz
Subcarrier spacing	3840 kHz
Channel BW (data subcarriers + guard bands)	2160 MHz
Number of BSs / numbers of UEs	1/1
Code Rates	0.75
SC-FDE role off	0.22
Modulation format	QPSK, 16-QAM, 64-QAM
PN compensation	CPE removal per DFT-s-OFDM symbol, or per SC-FDE block
Channel	AWGN
receiver type	linear (MMSE CE + MMSE equalization)
FEC	NR low-density parity-check (LDPC)
PA output power back-off dB	3 to 10 dB

The results of this comparison are shown in Figure 5-9. In this figure, only the block error rate (BLER) for an output back-off (OBO) of 10 dB is shown. It is obvious that in this case the performance is only limited by the phase noise. The comparison shows that there is a limited difference between SC-FDE and DFTS-OFDM with the given system configuration. However, it needs to be noted that in the shown simulation the subcarrier spacing for the DFTS-OFDM system is relatively large and the phase noise given by the model is not too large to limit system performance. Therefore, it is concluded that there is a benefit of using SC-FDE over DFTS-OFDM if PN is the limiting factor of the system. For the case that phase noise is not limiting system performance either SC-FDE or DFTS-OFDM can be chosen.

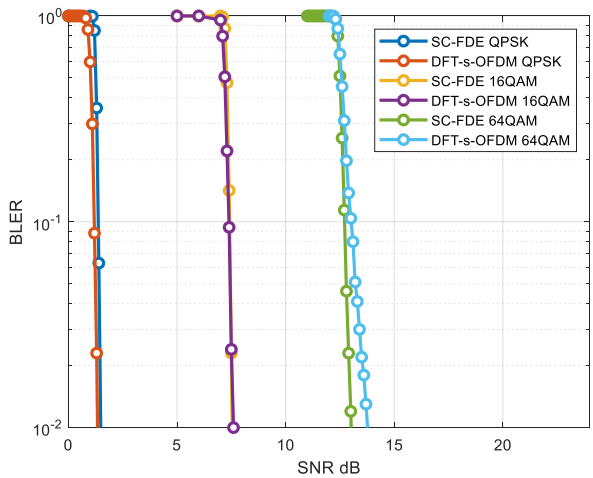


Figure 5-9: Comparison of SC-FDE and DFT-s-OFDM with phase noise and non-linear PA distortion.

5.1.2.1.2 Impact of PA nonlinearity

Beside evaluation in the phase noise limited regime also the PA non-linearity limit regime was evaluated. The results in Table 5-8 show the signal-to-noise ratio (SNR) in dB at which a BLER of 10^{-1} is reached. It is obvious that the SC-FDE can reach the same performance at between 1 – 2 dB less Output Backoff (OBO). Another interesting fact is that the performance at a high OBO of the DFT-s-OFDM system is slightly better than the SC-FDE system for QPSK and 16-QAM modulation. From these results, it can be concluded that given the PA model used in this work it is possible to use 1 – 2 dB less OBO for the SC-FDE system relative to the DFTS-OFDM system. As the PA output power will be a limiting factor of consumer grade communication systems operating in the frequency range from 100 to 300 GHz, which might justify the use of SC-FDE.

Table 5-8: SNR in dB at which a BLER of 10^{-1} is reached dependent on the OPBO.

OPBO in dB	DFTS-OFDM QPSK	SC-FDE QPSK	DFTS-OFDM 16-QAM	SC-FDE 16-QAM	DFTS-OFDM 64-QAM	SC-FDE 64-QAM
3	NA	4.12	NA	NA	NA	NA
4	7.00	2.39	NA	NA	NA	NA
5	3.17	1.92	NA	9.11	NA	NA
6	1.96	1.70	10.7	8.01	NA	14.5
7	1.55	1.55	8.20	7.72	18.1	13.4
8	1.36	1.46	7.67	7.57	13.8	13.0
9	1.25	1.40	7.49	7.47	13.2	12.8
10	1.19	1.37	7.39	7.42	13.0	12.7

5.1.2.1.3 Impact of combined phase noise and quantization noise

Besides the PAs at the transmit side, the analogue-to-digital converters (ADC) at the receive side may have a large part on the overall system power consumption, depending on the deployed network architecture i.e., on the number of radio frequency (RF) chains. ADC's power consumption doubles for every additional quantization bit. It is, therefore, desired to reduce the number of deployed quantization bits. On the other hand, reducing the number of quantization bits reduces ADC's resolution and may degrade the achievable system performance. It's obvious that there is an optimal point in the trade-off between the number of deployed quantization bits i.e., power consumption and achievable performance.

Simulative investigations on the effect of phase noise and quantization noise have been jointly carried out for SC-FDE and DFTS-OFDM transmissions. The purpose of the study is to gain an insight into the energy efficiency characteristics of these two waveforms.

Wideband transmissions of up to 4GHz and at the carrier frequency of 120 GHz with a subcarrier spacing of 960 kHz (for DFTS-OFDM) are considered. Further simulation parameter setting is summarized in Table 5-9.

Table 5-9: Parameter setting.

Parameter	Value
Carrier frequency (GHz)	120
FFT size	4096
Occ. BW (GHz)	O:3.17 / D:295 / S:3.92
Coding	LDPC 3/4
RRC β	0.3 (SC-FDE only)
Cyclic Prefix (%)	7
PT-RS (%)	3
Channel	AWGN
Nr.of Tx, Rx Antenna	1,1

Details on the applied PN model can be found in earlier report [HEX21-D22] section 3.6.6.2. PN contribution on the received signal is estimated and compensated according to simple mitigation algorithms for both waveforms (cf. [HEX21-D22] section 4.3.1). Uniform quantizer with equal spaced steps is used. For DFTS-OFDM a typical 4σ clipping is applied at the quantizer input where σ^2 is the variance of the zero-mean gaussian distribution of the received signal amplitude. For the block transmission SC-FDE signal the amplitude distribution is rather formed by the shape of the applied filter. An adaptive quantizer is therefore deployed here which dynamically adjusts the dynamic range and step size to the input amplitude on the symbol basis.

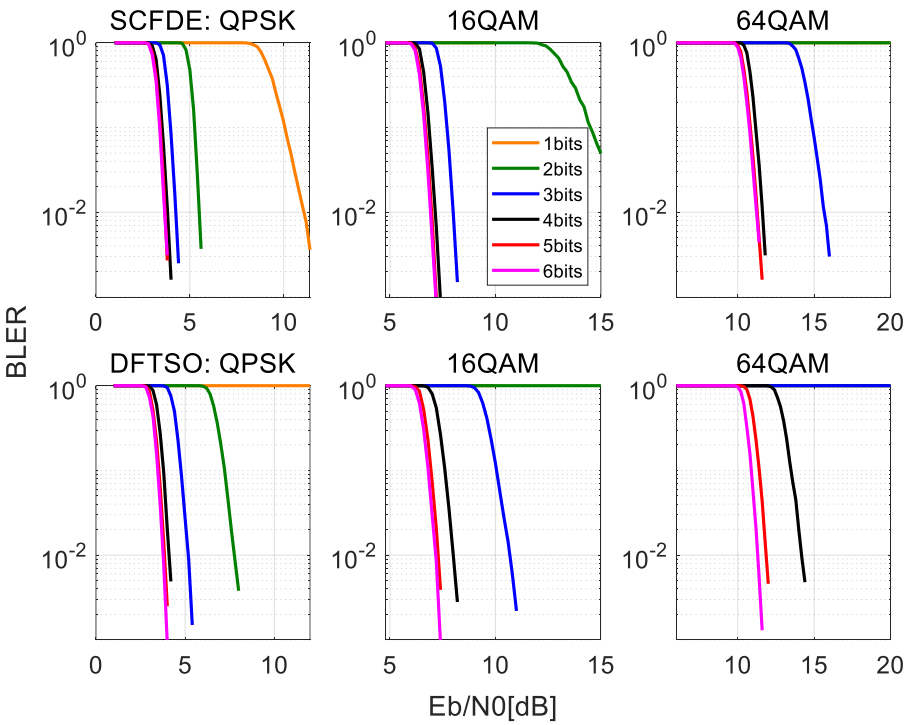


Figure 5-10: Dependency on the deployed number of quantization bits of SC-FDE (upper row) and DFT-S-OFDM (lower row) for QPSK, 16QAM and 64QAM.

Figure 5-10 shows the dependency of BLER performance on the number of deployed quantization bits of SC-FDE and DFTS-OFDM for QPSK, 16QAM and 64QAM. As can be seen from the plots, 4-5 quantization bits are in general sufficient for transmissions with 64QAM. Further SC-FDE shows a more robust behaviour in case of very coarse quantization. Compare, for example, the 64QAM case, optimal BLER performance can be reached with 4 quantization bits for SC-FDE transmission whereas 5 bits are needed for DFTS-OFDM to achieve the same BLER performance. Applying the power consumption model described in [Mur22] this means that the ADC power consumption using SC-FDE transmission waveform can be of a factor of 2 lower compared to the case of using DFTS-OFDM waveform.

5.1.2.2 Performance evaluation of reduced-PAPR SC-FDE and DFTS-OFDM waveforms

5.1.2.2.1 PA friendly constellation pattern and pulse shaping filter for SC-FDE

The signal amplitude envelope determines the peak to average power ratio (PAPR) level of the transmit signal. Thanks to the simple transmit structure of SC-FDE design the modulation alphabet can influence the envelope variation of the signal and hence reduce the PAPR. 3GPP specifies standard square QAM as modulation alphabets for NR, as exemplary shown for 16QAM in Figure 5-11 (left plot). It's primarily designed for the subcarrier of an OFDM signal, where the impact on PAPR from individual QAM schemes on the subcarrier is almost negligible. The impact of the QAM shape is relevant only for single carrier signals (and depends there also further on other parameters like filter shape). Another

standard utilizing single-carrier waveform is the digital video broadcasting second generation (DVB-S2) for satellite broadcasting [DVB-S2X]. As modulation constellation amplitude phase shift keying (APSK) patterns are defined. Looking at the left plot of Figure 5-11, it's intuitive that those corner constellation points on the outer ring contributes the most to the signal peaks. Rearranging them to other inner rings helps to reduce the variation in the signal envelope. On the other hand, reducing the Euclidian distances between the constellation points increases the decoding error probability at the receiver and hence the system performance. It is a trade-off between the PAPR reduction and the tolerable packing density of the constellation points.

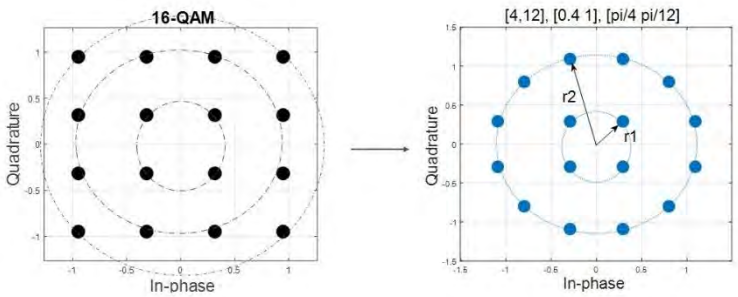


Figure 5-11: 4-bits constellation pattern for standard NR 16QAM (left) and PA friendly 16APSK (right).

A selected 2 rings constellation pattern is used now as an example [DVB-S2X]. The ratio of the ring radii, r_1/r_2 , is 0.6. The number of constellation points on the outer and inner rings is 12 and 4, respectively. Figure 5-12 shows the complementary cumulative density function (CCDF) distribution of the PAPR (left plot) and the BLER performance (right plot) of the chosen constellation pattern. As baseline PAPR and BLER performance of NR standard 16QAM constellation are used. Table 5-9 summarizes the main simulation parameters used to produce the results shown in Figure 5-12 & Figure 5-13.

It can clearly be seen from the plots that a large PAPR reduction of 1.8 dB can be achieved using 16APSK pattern compared to NR 16QAM case. It is, however, at cost of 0.3 dB loss in BLER performance. When considering PN with simple estimation and compensation method as described in [HEX21-D22], section 4.3.1, an additional 0.2 dB loss must be accounted for. In a scenario where PA output power is the limiting factor, the PAPR reduction can be well harvested for coverage extension. The loss in BLER performance can be compensated by e.g., a more robust channel coding rate i.e., at some spectral efficiency cost. Higher modulation orders tolerate smaller decoding error probability. Therefore, less PAPR reduction from reshaping constellation can be expected.

For the 6-bits modulation case, an exemplary APSK constellation with 8, 16, 20, 20 constellation points on 4 rings is chosen (inset of the left plot) [DVB-S2X]. A code rate of ¾ was used and the radian ratios are 0.27,0.59,0.97,1.40, respectively. As in case of 4-bits modulation above, Figure 5-13 depicts the CCDF distribution of PAPR (left) and BLER performance (right) for the 4-ring APSK modulation pattern (solid blue) and the baseline NR 64 QAM (solid black). It can be observed for the selected APSK constellation pattern, a PAPR reduction by 0.6dB can be achieved, without any further penalty in BLER performance for both cases, with and without PN consideration.

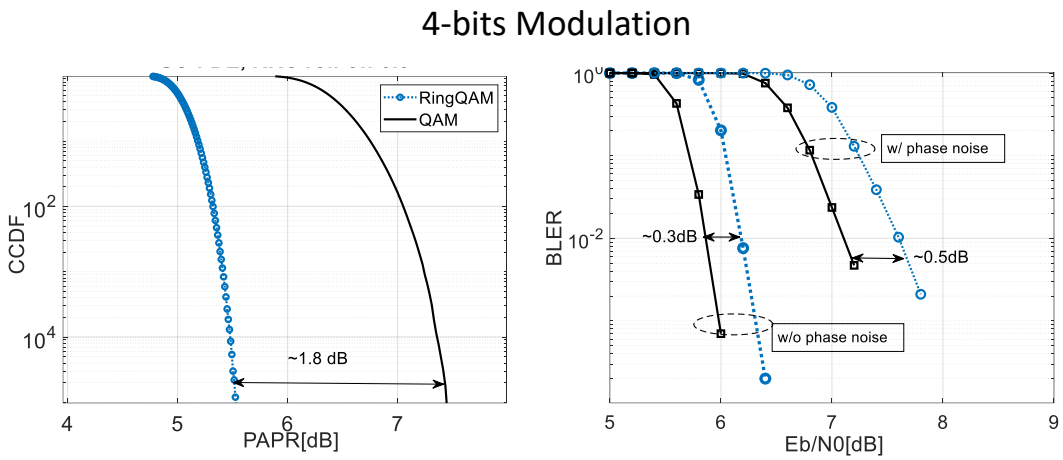


Figure 5-12: 4-bits modulation: CCDF of the PAPR (left) and BLER as a function of required Eb/No (right) of SC-FDE waveform using NR 16QAM and 16APSK constellation pattern.

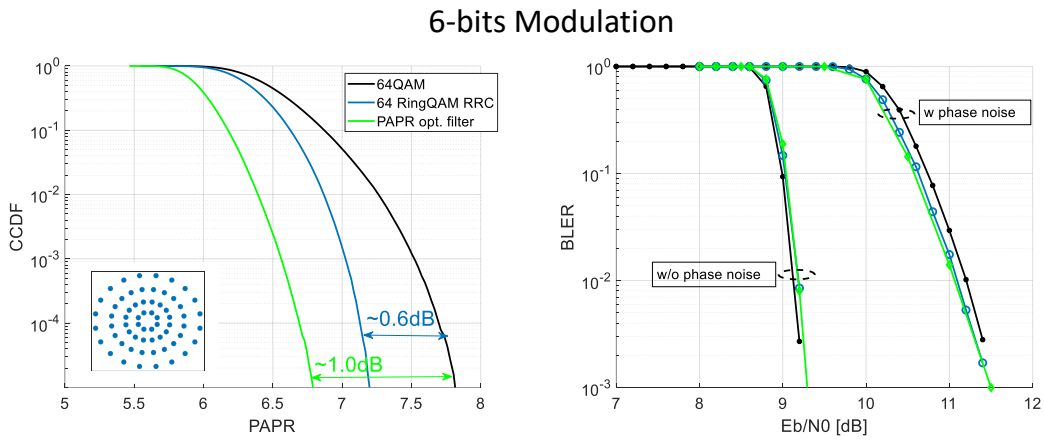


Figure 5-13: 6-bits modulation: CCDF of the PAPR (left) and BLER as a function of required Eb/No (right) of SC-FDE waveform using NR 16QAM and 16APSK constellation pattern. Inset: applied 4 rings constellation pattern for 64APSK.

As already mentioned above, pulse shaping filter is an additional design parameter for SC-FDE transmissions which can be used to constrain the signal envelope and hence is more PA friendly and efficiently extends the achievable transmission range. Standard root raised cosine (RRC) filter is commonly applied to band limit a transmission. Explicitly designing the pulse shaping filter to minimize the peak amplitude of the transmit signal allows to further reduce the PAPR level. The applied RRC filter shown in Figure 5-12 and Figure 5-13 above uses roll-off factor of 0.3. The designed PAPR aware filter, with similar excess bandwidth, achieves an additional 0.4 dB PAPR reduction, without further penalty in BLER performance (green curves in Figure 5-13).

The study in this subsection shows the potential of SC-FDE transmissions to reduce the PAPR level by selecting suitable constellation patterns and pulse shaping filter with moderate to no losses in the BLER performance. This PA friendly behaviour makes this waveform a potential candidate for transmissions in scenarios in which PA output power is a limiting factor and power/energy efficiency is of cardinal relevance.

5.1.2.2.2 DFTS-OFDM with and without frequency-domain filtering

With the operation of mobile networks moving to sub-THz carrier frequencies, a number of hardware-related challenges arises. One such challenge is the decreasing peak power and power efficiency of PAs, the reason for which is the shrinking of integrated circuit geometries with increased operating

frequency, which in turn requires lower operational power to avoid dielectric breakdown. Decreased power capability of PAs will naturally compromise coverage/performance at sub-THz, already compromised by increased pathloss. One way of improving coverage consists of pushing the operating point of PAs closer to saturation power level without increasing the level of nonlinear distortion, which may be achieved by choosing waveforms with small envelope variations. One such waveform is DFT-spread OFDM (DFTS-OFDM), which is akin to ordinary OFDM but with a smaller envelope variation. In this contribution, ordinary DFTS-OFDM and a modification of DFTS-OFDM (employing frequency-domain filtering) are compared in terms of potential coverage extension, under the influence of a nonlinear PA.

DFTS-OFDM and FDF-DFTS-OFDM

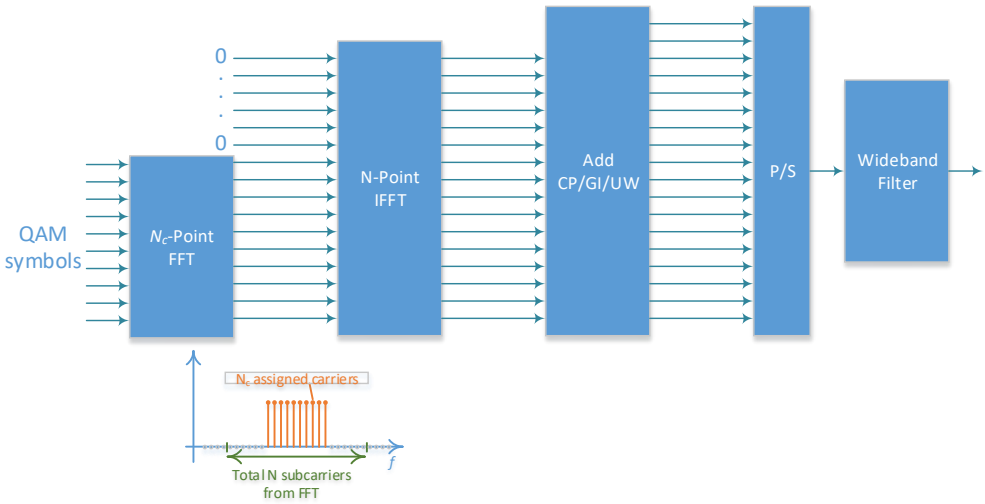


Figure 5-14: Complex baseband of a DFTS-OFDM transmitter.

DFTS-OFDM waveform is synthesized by performing a DFT of the vector of modulated symbols (modulated e.g. using QAM) before mapping them to OFDM subcarriers, cf. Figure 5-14.

Another waveform candidate analyzed in this contribution is DFTS-OFDM with Nyquist/square root Nyquist filtering in frequency domain, referred to in the follow-up as frequency-domain filtered DFTS-OFDM (FDF-DFTS-OFDM). There are two versions of FDF-DFTS-OFDM: one with bandwidth expansion, and the other without bandwidth expansion.

The principle of operation of FDF-DFTS-OFDM with bandwidth expansion is described and illustrated in Figure 5-15. The application of a square root Nyquist filter ensures that there is no additional frequency selectivity after combining and hence no penalty in performance. Additionally, frequency-domain filtering with a non-zero rolloff β will serve to smoothen the envelope variations of the time domain signal, and therefore result in smaller envelope variations compared to ordinary DFTS-OFDM. If the bandwidth used by the N_c data-bearing subcarriers in Figure 5-15 is BW_{data}, total occupied bandwidth after replicating and pulse shaping is

$$BW_{occupied} = (1+\beta)BW_{data}. \quad \textbf{(5-1)}$$

This bandwidth expansion is illustrated by the grey-coloured "tails" of the spectrum at the IFFT input in Figure 5-15. On the other hand, if the occupied bandwidth is constrained to be exactly BW_{data} (e.g. in order to satisfy coexistence requirements), then the number of data-bearing subcarriers has to be reduced so that they occupy a reduced bandwidth

$$BW'_{data} = \frac{BW_{data}}{(1+\beta)}. \quad \textbf{(5-2)}$$

This case is denoted as FDF-DFTS-OFDM without bandwidth expansion. In order to ameliorate the reduction of transmit power caused by reduced bandwidth, the pre-shaping signal power should be boosted by a factor of $1+\beta$ (this will be referred to as power boosting in the follow-up). Moreover, the

loss of raw throughput induced by bandwidth reduction (2) can be compensated by increasing the code rate by a factor of $1+\beta$. The principles of power boosting and of bandwidth expansion/no bandwidth expansion are illustrated in Figure 5-16, where $(1-\gamma) = 1/(1+\beta)$.

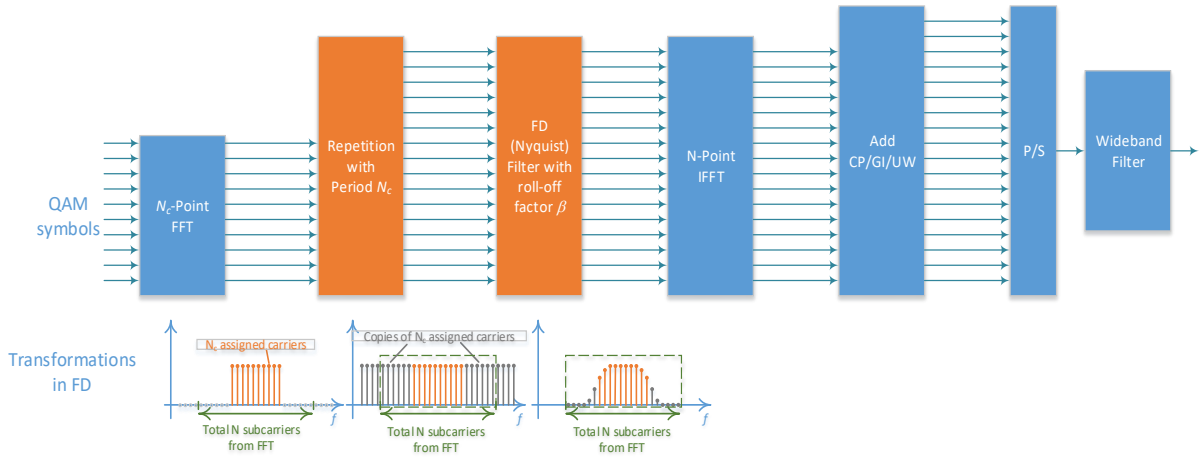


Figure 5-15: Complex baseband of a FDF-DFTS-OFDM transmitter.

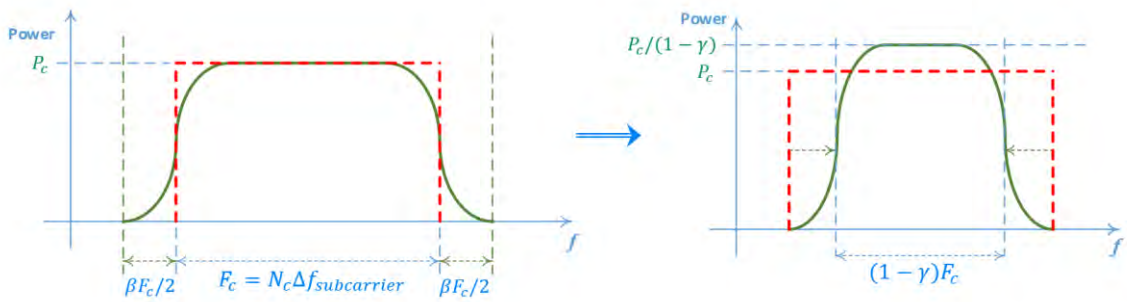


Figure 5-16: Principle of power boosting.

If power boosting and code rate increase are performed, a system using FDF-DFTS-OFDM without bandwidth expansion thus has the same power and throughput as a system using an ordinary DFTS-OFDM signal, which allows for a fair comparison with other methods in terms of transmit power (consequently, SNR) and occupied bandwidth. Therefore, FDF-DFTS-OFDM without BW expansion will be a matter of study in this contribution.

A comparison of candidate waveforms in terms of PAPR is shown in Figure 5-17 for a set of standard modulation schemes. In addition to ordinary pi/2 – BPSK for DFTS-OFDM, a time-domain filtered version of pi/2 – BPSK modulation with discrete filter impulse response [1 -0.28] is analyzed – an alternative way of reducing the envelope variations of DFTS-OFDM that introduces frequency selectivity and hence performance deterioration.

Subcarrier spacing (SCS) is 960 kHz and the bandwidth is 1.88 GHz for all waveforms and modulations. PAPR statistics are measured on a block level, where one PAPR sample is measured over 13 OFDM/DFTS-OFDM/FDF-DFTS-OFDM symbols carrying only data modulated by the modulations indicated. One can observe how FDF-DFTS-OFDM provides a clear reduction of PAPR compared to DFTS-OFDM (1.5 – 4 dB depending on the modulation).

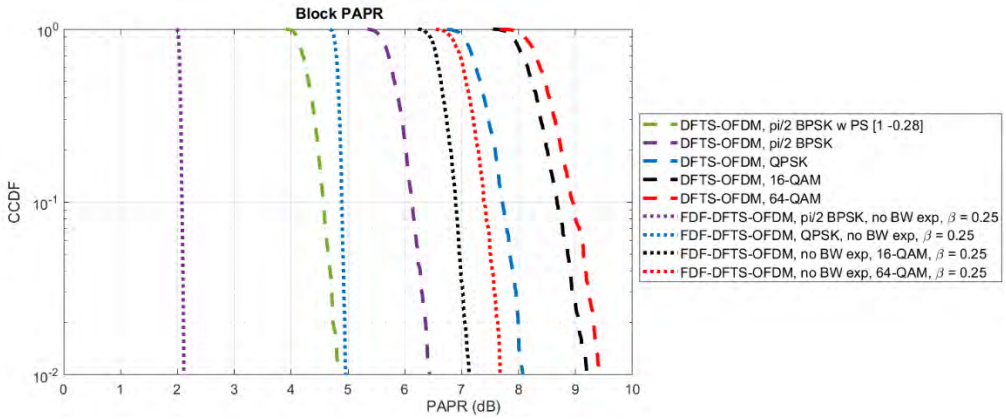


Figure 5-17: Block PAPR for the set of chosen waveforms.

Link budget – based comparison of DFTS-OFDM and FDF-DFTS-OFDM

Measuring PAPR of a waveform gives a good measure of envelope variability, but does not directly indicate the link budget improvement (and correspondingly, coverage extension) relative to another waveform used. In order to correctly assess the link budget improvement, following steps need to be undertaken:

1. Calculating the minimum OBO of the PA when using the waveform, given a set of error vector magnitude (EVM) and out-of-band (OOB) radiation constraints;
2. Calculating the minimum Rx power needed to support a certain performance (e.g. a certain throughput)
3. Combining 1 and 2 in a link budget calculation to assess the amount of link budget headroom used to cover propagation-related power losses. The calculated link budget headroom translates directly to propagation distance (coverage).

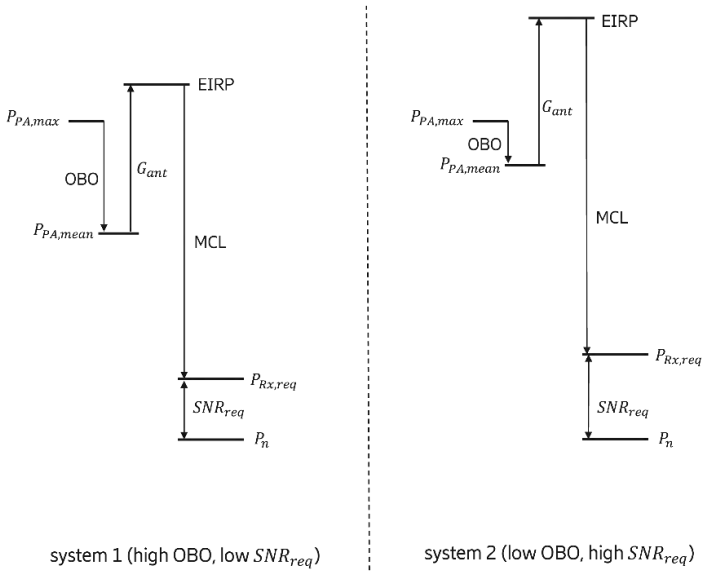


Figure 5-18: link budget comparison of two systems using different waveforms.

A companion illustration for the proposed system comparison methodology is given in Figure 5-18. The link budget on the left-hand side corresponds to a hypothetical system using a waveform with high envelope variation, requiring a large PA OBO to ensure low probability of clipping and consequently low distortion. The link budget on the right-hand side may correspond to a system using a waveform with low envelope variation but requiring a higher demodulation SNR to support the same performance as system 1. There is a difference in the link budget headroom (labelled as maximum coupling loss

(MCL) in the figure) between the two systems, with system 2 having a larger MCL and therefore being able to accommodate a larger propagation loss.

DFTS-OFDM vs FDF-DFTS-OFDM: comparison of minimum power backoff

Minimum power backoff of the PA is the difference (in dB) between the saturation power of the PA and the mean output power, chosen to ensure that in-band and out-of-band performance requirements are satisfied. In-band requirements are quantified by error vector magnitude (EVM) requirements, given in Table 5-10.

Table 5-10: EVM requirements.

Modulation	EVM limit value
$\pi/2$ – BPSK	30 %
QPSK	17.5 %
16-QAM	12.5 %
64-QAM	8 %

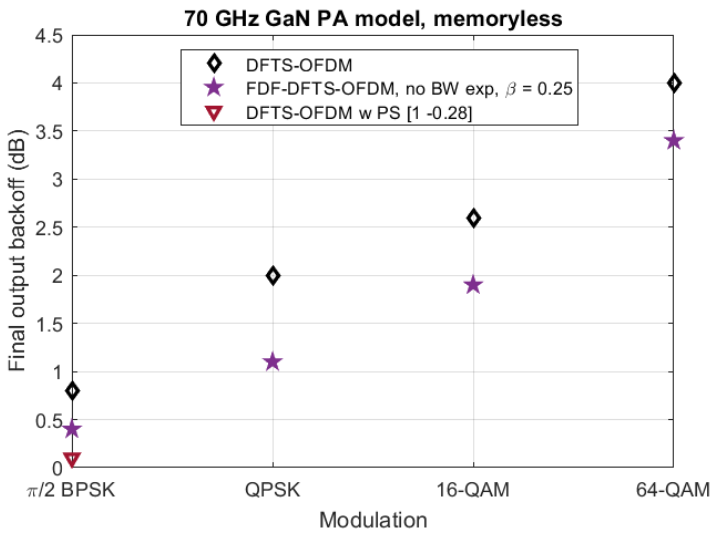


Figure 5-19: Minimum OBO satisfying EVM and ACLR requirements.

OOB requirements are quantified by adjacent channel leakage ratio (ACLR), calculated as the ratio of signal power in configured bandwidth BW_{config} centered at the carrier frequency f_c (in-band power), and signal power in configured bandwidth BW_{config} centered at frequency $f_c + BW_{channel}$, where $BW_{channel} \geq BW_{config}$. ACLR requirement for the evaluations in this contribution is set to 26 dB. The PA model used in the evaluations is the memoryless GaN model described in [HEX21-D22, Section 3.6.6.1]. As for other parameters of the evaluation, subcarrier spacing of both waveforms is set to 960 kHz, $BW_{channel} = 2.16$ GHz and $BW_{config} = 1.88$ GHz. Results are shown in Figure 5-19, and indicate that use of frequency-domain shaping reduces the OBO by 0.5 – 1 dB, depending on the modulation used.

DFTS-OFDM vs FDF-DFTS-OFDM: link level performance

In the second step in the analysis, link-level performance of the two waveforms under investigation is assessed. The set of modulation and coding schemes (MCSs) selected for the analysis is taken from [38.214, Table 6.1.4.1 – 1, q = 1]. The MCSs for FDF-DFTS-OFDM are the same as for DFTS-OFDM but with code rate adjusted:

$$r_{c,FDF-DFTS-OFDM} = r_{c,DFTS-OFDM}(1 + \beta). \quad \textbf{(5-3)}$$

A selection of link level performance results is shown in Figure 5-20, illustrating the differences in BLER performance between DFTS-OFDM and FDF-DFTS-OFDM. MCSs selected for FDF-DFTS-OFDM match those for DFTS-OFDM in throughput, as they effectively use less BW. For the evaluations, practical channel estimation was used. Channel model is SISO CDL/TDL-A with delay spread of 10 ns and UE speed of 3 km/h. Block fading was assumed in the evaluations, i.e. a random channel impulse response is generated for each block, with block equal to one slot (14 DFTS-OFDM symbols).

The results clearly show the negative impact on performance that the code rate increase at higher modulation orders has on the performance of FDF-DFTS-OFDM, the effect which will eventually offset the OBO gains when considered in the context of full link budget evaluation.

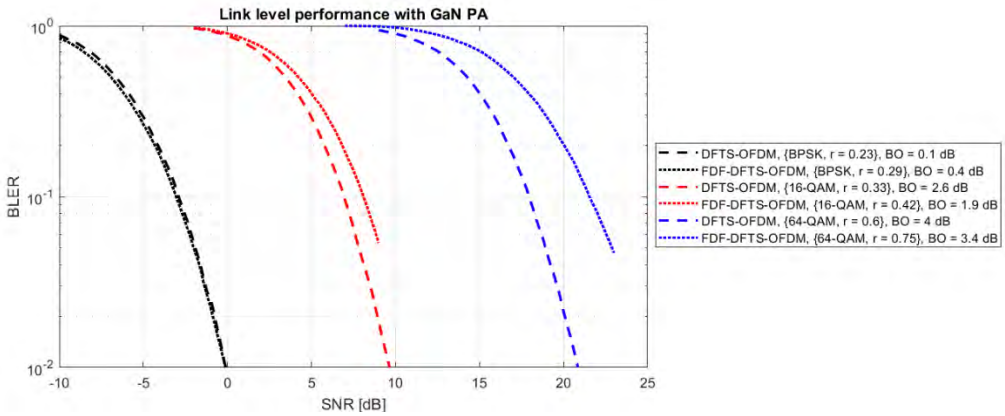


Figure 5-20: BLER comparison of DFTS-OFDM and FDF-DFTS-OFDM under the influence of PA nonlinearity.

DFTS-OFDM vs FDF-DFTS-OFDM: total link budget comparison

Motivated by the illustration in Figure 5-18, let us define the maximum coupling loss (MCL) as

$$MCL = P_{PA,max} - OBO + G_{ant} - (SNR + P_n)\ [dB], \quad \textbf{(5-4)}$$

where P_n is thermal noise power and G_{ant} total antenna and array gain at Tx and Rx side. Let us also define the information throughput as

$$Th = SE(1 - BLER)\left[\frac{bps}{Hz}\right], \quad \textbf{(5-5)}$$

where SE denotes spectral efficiency. Now, between two transmission systems with same throughput but different MCL supporting those throughputs, the system having a larger MCL achieves larger coverage. Conversely, when comparing two systems operating at the same MCL, the system achieving a larger throughput is "better" since it achieves a better throughput at the same coverage as the other system.

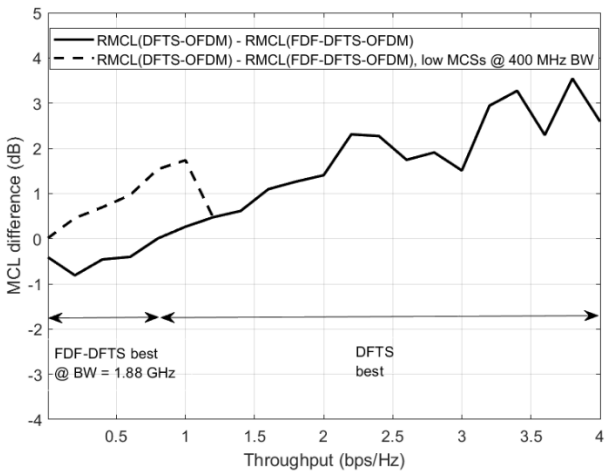


Figure 5-21: MCL difference vs throughput.

For the analysis in this section, the envelope of throughput curves is calculated for all MCSs simulated, and the obtained throughput is plotted against the difference of MCL for a particular throughput in Figure 5-21. The described 3-step analysis is repeated also for a narrow bandwidth allocation of 400 MHz for lower order modulations and those results are also shown.

What the results reveal is that FDF-DFTS-OFDM has advantage over DFTS-OFDM only at the lowest spectral efficiencies, and that it provides at most 0.8 dB increase in MCL (corresponding to a coverage extension of about 10% if free-space pathloss is assumed). Therefore, FDF-DFTS-OFDM may be considered for supporting a moderate coverage extension for the users at the cell edge, whereas for other users DFTS-OFDM is preferred.

5.1.2.3 Optimal pilot configuration for phase noise compensation at 140 GHz

This section presents a systematic numerical analysis on the pilot design for phase noise compensation of DFTS-OFDM waveform at the 140 GHz band. The pilot scheme used in this analysis follow the approach of the phase tracking reference signals (PTRS) of 3GPP [38.211], which consists of 2, 4 or 8 pilot groups with samples each. In addition, in order to have a more general set of pilots which is more adequate to the sub-THz channel, an extension is considered with 16, 32 and 64 groups of pilots to analyse the performance and data rate trade-off. In addition, each pilot group occupies 4 time-domain samples to have a comparable setup with the PTRS of 3GPP. For compactness, only the indexes of the first pilot of each group is displayed in Table 5-11 for N_g =16, 32 and 64 groups, which were defined by distributing the pilots as equally as possible along the time domain samples. For example, the pilots for $N_g = 16$ are [**1** 2 3 4 **137** 138 139 140 …], where the indexes [1 137 …] represent the first index of each pilot group.

Table 5-11: Indices of the time-domain pilots for PN estimation

N_g	Indexes of first pilot of each pilot group
16	[1, 137, 274, 411, 548, 685, 821, 957, 1093, 1229, 1365, 1501, 1637, 1773, 1909, 2045]
32	[1, 67, 133, 199, 265, 331, 397, 463, 529, 595, 661, 727, 792, 857, 923, 989 1055 1121 1187, 1253, 1319, 1385, 1451, 1517, 1583, 1649, 1715, 1781, 1847, 1913, 1979, 2045]
64	[1, 33, 65, 97, 129, 161, 193, 225, 257, 289, 321, 353, 385, 417, 450, 483, 516, 549, 582, 615, 648, 681, 714, 747, 780, 813, 846, 879, 912, 945, 978, 1011, 1044, 1077, 1110, 1143, 1176, 1209, 1242, 1275, 1308, 1341, 1373, 1405, 1437, 1469, 1501, 1533, 1565, 1597, 1629, 1661, 1693, 1725, 1757, 1789, 1821, 1853, 1885, 1917, 1949, 1981, 2013, 2045]

The phase noise compensation algorithm estimates the average phase noise coefficient per pilot group and performs a linear interpolation to estimate the phase noise rotation at each time-domain symbol. The phase noise model is based on the 3GPP document [HEX21-D22, Section 3.6.6.2]. This model has been verified up to the frequency of 70 GHz. In this work, a frequency f1 = 70 GHz is generated by a phase locked loop (PLL) and then multiplied to generate f2 = 140 GHz. This process will result in a phase noise with similar power spectral density (PSD) as the 70 GHz system but with 6 dB more power. In order to verify solely the effect of each pilot scheme, the underlying channel model is additive white gaussian noise (AWGN) which in practice corresponds to a highly directional antenna where multipath components are too small.

The goal of the simulation shown in this section is to evaluate the performance loss of the DFTS-OFDM system at 140 GHz due to phase noise. This analysis in terms of Eb/N0, where the pilot overhead is considered. In particular, the goal is to compare the performance loss for given SCS and pilot scheme such that recommendations can be given for each scenario. The phase noise effects of both BS and UE sides are simultaneously considered. Regarding the simulation parameters, an FFT size of 2048 is considered with SCS of 960, 1920 and 3840 kHz. For the coding scheme, the NR-based LDPC coding with R = ¾ coding rate is considered. Regarding the pilot configuration, $N_g = 2, 4, 8, 16, 32$ and 64 time-domain pilot groups are considered with 4 time-domain pilots in each group, where $N_g = 2, 4, 8$ consider PTRS of 3GPP [38.211], and $N_g = 16, 32, 64$ are extensions with equally spaced pilot groups with pilots given in Table 5-11. Lastly, the BLER results with no phase noise is shown as a benchmark.

Figure 5-22 first shows the BLER results for 16-QAM. A summary of the performance loss at BLER = 10^{-2} of the system under phase noise is comparison to the phase noise free system is given in Table 5-12. The combination of SCS and N_g that lead to interesting practical options are highlighted. Firstly, it can be observed that the pilot configuration with $N_g = 32$ achieves the best performance trade-off in

terms of energy per bit with the minimum performance loss of 0.9, 0.9 and 0.8 dB for the SCS of 960, 1920 and 3840 kHz, respectively. Increasing the number of pilot groups to $N_g = 64$ increases the performance loss to 1 dB because the amount of energy spent in the additional pilots does not compensate the small performance gain in terms of SNR, which does not make it an interesting choice. In addition, it is important to highlight that the pilot group number of $N_g = 16$ is also an option to be used, especially with SCS of 1920 and 3840 kHz, since the performance loss in relation to $N_g = 32$ is small. For SCS of 3840 kHz, the configuration $N_g = 8$ is an option for the same reason.

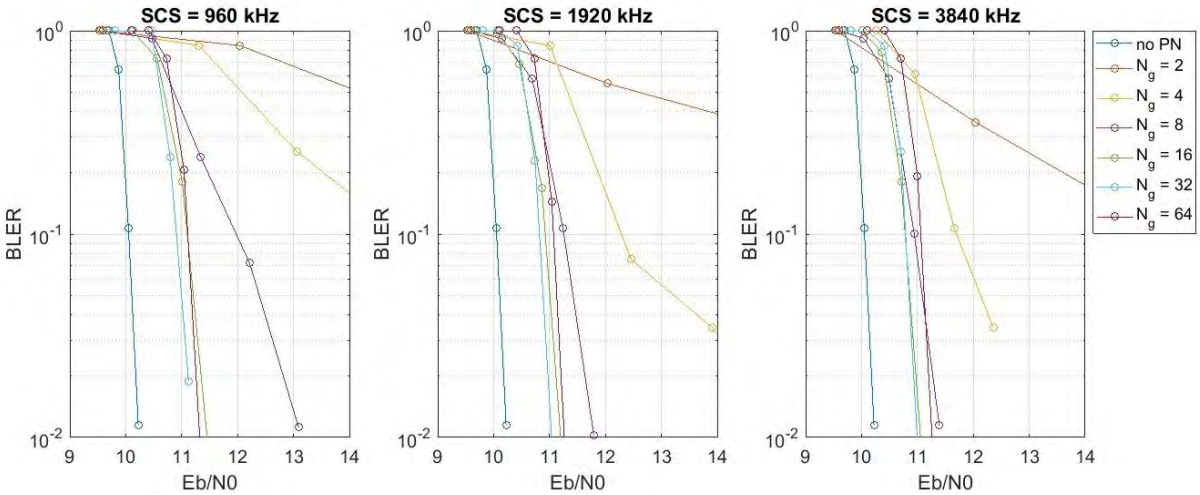


Figure 5-22: Block Error Rate (BLER) for 16-QAM, R = ¾, subcarrier spacing of 960, 1920 and 3840 kHz, with N_g=2,4,8,16,32 and 64 pilot groups.

Table 5-12: Loss in terms of Eb/N0 for 16QAM pilot schemes at BLER = 10^{-2}. NA = not achieved.

SCS	$N_g = 2$	$N_g = 4$	$N_g = 8$	$N_g = 16$	$N_g = 32$	$N_g = 64$
960 kHz	NA	NA	2.9 dB	1.2 dB	0.9 dB	1dB
1920 kHz	NA	5.1 dB	1.6 dB	1 dB	0.8 dB	1dB
3840 kHz	NA	2.7 dB	1.2 dB	0.85 dB	0.8 dB	1dB

The simulation has also been carried out for 64-QAM with 3/4 code rate NR LDPC code. The performance loss results are shown in Table 5-13. As expected, for the higher MCS, the phase noise has a higher impact in the performance loss. In addition, one important aspect to notice is that for SCS of 960 kHz, all the PTRS of 3GPP [38.211] with of $N_g = 2$, 4 and 8 pilot groups do not achieve the BLER of 10^{-2}, meaning that the extended set of pilots is necessary in this case. Also, the system with $N_g = 64$ pilot groups provide the best trade-off for SCS = 960 and 1920 kHz, however, it is important to notice that the performance gain of $N_g = 64$ in relation to $N_g = 32$ is relatively small, e.g., 0.4 and 0.1 for the SCS of 960 and 1920 respectively. For SCS = 3840 kHz, there is almost no gain of $N_g = 64$ in relation to $N_g = 32$. Thus, in practise $N_g = 32$ can be a suitable choice because it leads to less pilot overhead. Lastly, for the SCS of 3840 kHz, $N_g = 16$ is also an option due to small loss with higher N_g.

Table 5-13: Loss in terms of Eb/N0 for 64QAM pilot schemes at BLER = 10^{-2}. NA = not achieved.

SCS	$N_g = 4$	$N_g = 8$	$N_g = 16$	$N_g = 32$	$N_g = 64$
960 kHz	NA	NA	3.2 dB	1.5 dB	1.1dB
1920 kHz	NA	NA	1.7 dB	1 dB	0.9 dB
3840 kHz	NA dB	2.3 dB	1.2 dB	0.9 dB	0.9 dB

Although these simulations considered 16-QAM and 64-QAM for FFT size of 2048, comments about possible outcomes for other configurations can be discussed. In particular, it can be expected that the FFT size does not play an important role in this case because it does not change the time duration of the DFT-s symbol, which is solely dependent on the SCS choice. Also, as the MCS gets larger, e.g., 128/256 QAM, the required N_g increases as observed from 16-QAM to 64-QAM. In addition, another case is considered using a highly directive antenna where the multipath components (MPCs) are too small such

that the channel collapses into an AWGN. In general, if there are MPCs, it is expected that an extra performance loss due to imperfect channel estimation and additional noise in the equalization process at the receiver, i.e., the necessary values of N_g found in Table 5-12 and Table 5-13 are expected to increase, which increases the need of using the pilot schemes defined in Table 5-11.

5.1.3 Zero-crossing modulation

5.1.3.1 Introduction of Zero-Crossing Modulation

Zero-crossing modulation (ZXM) [FDB+19] is a modulation scheme that was designed under consideration of the hardware constraints imposed by ADCs. The power consumption of ADCs starts to increase quadratically with the sampling frequency for bandwidths larger than 300 MHz [Mur22], as envisioned for 6G. For systems with large antenna arrays and many RF chains, ADCs will therefore likely be a major contributor to the overall power consumption of the system [NSJ+21]. As the amount of power consumed by ADCs increases exponentially with its number of bits, ZXM is based on the idea, that sampling at high frequencies may still be feasible if the ADC complexity, i.e., its amplitude resolution, is reduced. In the extreme case of 1-bit quantization, receiver designs without the need of an automatic gain control (AGC) and with reduced linearity requirements for the analogue frontend might be feasible. With 1-bit ADCs the receiver has only access to the sign of the received signal and can effectively only determine the position of zero-crossings in time. Therefore, ZXM aims to encode the information to be conveyed in the distances of zero-crossings and the achievable rate therefore depends on the temporal grid on which zero-crossings can be placed.

A possible mean to create a ZXM modulated transmit signal is to combine faster-than-Nyquist (FTN) signalling and runlength-limited (RLL) sequences [LDF15]. RLL sequences [Imm04] are bipolar sequences that can be constructed from so called (d,k)-sequences. The latter are binary sequences with two constraints: Every 1 has to be followed by at least d 0s and at most by k 0s. To maximize the entropy rate of the transmit signal, the k-constraint is omitted, i.e., $k = \infty$. To construct an RLL sequence from a (d,k)-sequence, non-return-to-zero inverse (NRZI) encoding is used. This is illustrated in the following example for $d = 1$:

$$(d,k) \quad \bar{\boldsymbol{a}}_m = [\dots,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 1,\dots]^T$$

$$\text{RLL} \quad \boldsymbol{a}_m = [\dots,-1,\ 1,\ 1,-1,-1,-1,\ 1,\dots]^T$$

The complex valued transmit symbols are constructed from two real valued, independently encoded RLL sequences and are linearly modulated onto a waveform using a real valued transmit filter. Note that RLL codes have a code rate $R \leq 1$, i.e., encoding independent and identically distributed (i.i.d.) bits into an RLL sequence reduces the entropy per symbol to less than 1 bit. However, when combining RLL sequences with FTN signalling, i.e., transmitting more than one symbol per Nyquist interval, the entropy rate of the transmit signal can be increased to more than 1 bit/Hz/s. Note that FTN signaling comes at the cost of self-introduced intersymbol interference (ISI). However, the ISI can be well controlled by the minimum runlength of the RLL code, such that the oversampled receive signal allows for efficient RLL decoding [FDB+19].

5.1.3.2 Phase Noise Model

Another challenge that has to be dealt with in communications systems utilizing large bandwidths stems from the fact that typically large bandwidths imply high carrier frequencies, due to the availability of spectrum in the mmWave and sub-THz range. Common oscillator PN models, including the one presented in 4.2, show a quadratic dependency between the PN power and the carrier frequency, thus, high carrier frequencies necessitate the tracking and compensation of phase noise [YMY+19]. Phase noise estimation and compensation is a well-researched topic for systems with high resolution ADCs, however, there is very little research concerning the estimation of phase noise in systems with coarse quantization.

For this investigation, the phase noise model from [KKP+14] is adapted, which models the phase noise as a non-stationary process and that is commonly used to describe the phase noise behaviour of free running oscillators. Here, the phase noise stems from two independent processes:

$$\boldsymbol{\phi_k} = \boldsymbol{\phi_{0,k}} + \boldsymbol{\phi_{2,k}},$$

where $\phi_{0,k}$ is a white Gaussian phase noise process with variance $\sigma_{\phi_0}^2$ and $\phi_{2,k}$ is cumulative phase noise processes, that stems from the integration of white noise inside the oscillators and can be expressed in terms of an initial phase and a sum of phase increments:

$$\phi_{2,k} = \phi_{2,1} + \sum_{i=1}^{k} \xi_{2,i}$$

It can be shown that the phase increments are Gaussian distributed with variance σ_ξ^2. Moreover, the increments $\xi_{2,i}$ have no temporal correlation, i.e., $\phi_{2,k}$ is a Wiener process. In [KKP+14], there is an additional third phase noise component, $\phi_{3,k}$, that is left out here, as its energy is mostly concentrated at very low offset frequencies, and which therefore can almost be treated as a constant phase offset.

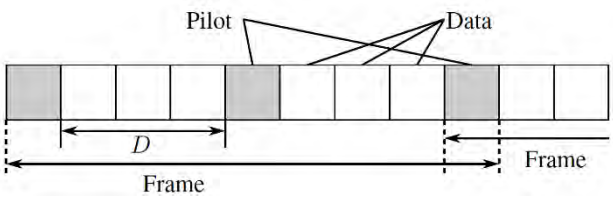


Figure 5-23: Model of the time frame structure, taken from [GSD+22].

5.1.3.3 Phase Noise Estimation

In this section, the estimation of phase noise in systems with 1-bit quantization and Nyquist-rate signalling and symbol rate sampling is investigated. Two practical estimation approaches and a theoretical lower bound on the estimation error are considered. Moreover, only pilot symbol aided phase noise estimation is employed, meaning that the sequence of transmit symbols periodically contains blocks of pilot symbols, whose values and temporal positions are known by the receiver a-priori. These considerations lead to the frame structure that is depicted in. All frames start and end with a pilot block and two pilot blocks are separated by D data blocks. Note that the data- and pilot blocks are assumed to be of equal length, consisting of N symbols. As a causal algorithm and a non-causal algorithm for phase noise estimation are used, it is required to define a finite frame for the non-causal one.

The estimation algorithms, first presented in [GSD+22], assume that the phase noise process varies slowly w.r.t. the symbol rate: for the derivation of the estimation algorithms, the phase noise has been modelled as block wise constant to simplify their structure. The block size was chosen to be equal to the pilot block size N, and only one estimate per block is produced. This leads to a block-based tracking approach, which is common practice [MMF97, Sec. 5.2]. The algorithms are two stage estimators, where the first stage consists of the constant phase LS estimator derived in [SDF22], working on the individual pilot blocks. The estimate of the LS estimator is given by

$$\hat{\phi}_{LS}(m) = \arg\left(\sum_{k=mN+1}^{(m+1)N} x_k^* y_k\right), \tag{5-6}$$

where x_k and y_k denote the samples from the noise free receive signal and the receive 1-bit quantized samples, respectively. The index of the current pilot block within the observation frame is denoted by m. This estimate can be interpreted as an estimate of the average phase within the pilot block, i.e.,

$$\hat{\phi}_{LS}(m) = \bar{\phi}(m) = \frac{1}{L_p} \sum_{k=mN+1}^{(m+1)N} (\phi_{0,k} + \phi_{2,k}). \tag{5-7}$$

In the post-processing stage, using knowledge on the statistics of the phase noise process, the estimates of the pilot blocks are interpolated to acquire phase estimates for the data blocks. Here, two different algorithms are used: a Kalman filter and the Rauch-Tung-Striebel (RTS) algorithm. The Kalman filter works in two steps: a model-based prediction step is followed by a data-based correction step. It is capable of using all observations from the past and is therefore a causal filter. The RTS algorithm forms a non-causal extension to the Kalman filter and therefore requires a closed observation frame. It performs a forward and a backward filtering sweep that both use the idea of the Kalman filter, but in the backward sweep on the time-reversed sequence. Note that the latter approach leads to a lower MSE at the cost of additional latency.

To set the results of the estimation algorithms into perspective, the optimal estimation performance that can be achieved based on 1-bit quantized observations was also examined in [ZGD+22]. For this purpose, the Bayesian Cramér-Rao bound (BCRB) was formulated. The BCRB is a lower bound for the achievable Bayesian mean-squared error (MSE) of any estimator and is given by the inverse of the so-called Bayesian information matrix that can be formulated as a sum of the expected Fisher information and a term that reflects the a-priori knowledge that estimators can employ:

$$\boldsymbol{B} = \mathrm{E}_{\boldsymbol{\phi}}[\boldsymbol{F}(\boldsymbol{\phi})] + \mathrm{E}_{\boldsymbol{\phi}}\left[-\nabla_{\boldsymbol{\phi}}^2 \log p_{\boldsymbol{\phi}}(\boldsymbol{\phi})\right]. \tag{5-8}$$

In the equation above, $\boldsymbol{\phi}$ denotes the vector of estimation parameters, which in this case corresponds the phase noise samples. Calculating the matrix for the system model presented here is not straightforward and the details of the derivation can be found in [ZGD+22]. It is important to note that the bound has been calculated only holds for estimators that create estimates solely based on observations made during pilot observations and discard available information about the estimation parameters that is contained in the observations that are made during data transmission. However, this covers most practical estimation algorithms, as it allows the separation of synchronization and data detection and, thus, enables low complexity receiver algorithms. The calculated bound is given by

$$\boldsymbol{B}^{-1} = \left(\frac{c_1 E_s}{\pi \sigma_n^2} e^{-\frac{c_2 E_s}{\sigma_n^2}} \left(I_0\left(\frac{c_2 E_s}{\sigma_n^2}\right) + I_1\left(\frac{c_2 E_s}{\sigma_n^2}\right)\right) \boldsymbol{\Delta}_P + \boldsymbol{C}_{\boldsymbol{\phi}}^{-1}\right)^{-1}, \tag{5-9}$$

where σ_n^2 is the variance of the additive noise in the channel, E_s denotes the symbol energy (which for this formulation is assumed to be constant for all symbols), $c_1 \approx 4.0360$ and $c_2 \approx 0.3930$ are numerically determined approximation constants, and $I_n(x)$ denotes the modified Bessel function of the first kind and n-th order. The pilot-data structure is captured in the matrix $\boldsymbol{\Delta}_P$ that is a diagonal matrix with ones on the main diagonal where pilot blocks are located within the observation frame and zeros where data blocks are located. The matrix $\boldsymbol{C}_{\boldsymbol{\phi}}$ denotes the covariance matrix of the estimation parameter vector $\boldsymbol{\phi}$. As the RTS algorithm, the BCRB uses all available information that can be extracted from the temporal correlation of the phase noise process and therefore, also requires a finite observation frame.

In Figure 5-24 the Bayesian MSE of the estimation algorithms is shown, along with BCRB, i.e., the diagonal elements of $\boldsymbol{B}^{-1}$, over the time index k. The parameters used for this evaluation can be found in Table 5-14. Naturally, the bound and the algorithms show the lowest MSE at the positions of the pilot blocks, which explains the periodic minima in the MSE curves. These minima are very pronounced for the Kalman filter but can be barely seen for the RTS algorithm and the BCRB, indicating that the RTS algorithm could be used with even larger pilot block spacing. Moreover, the RTS algorithm performs very close to the BCRB, and the MSE of the RTS algorithm and the BCRB exhibit a “bathtub” shape. This is intuitive, as the missing observations before and after the observation frame explain the rise of the MSE curve at the ends.

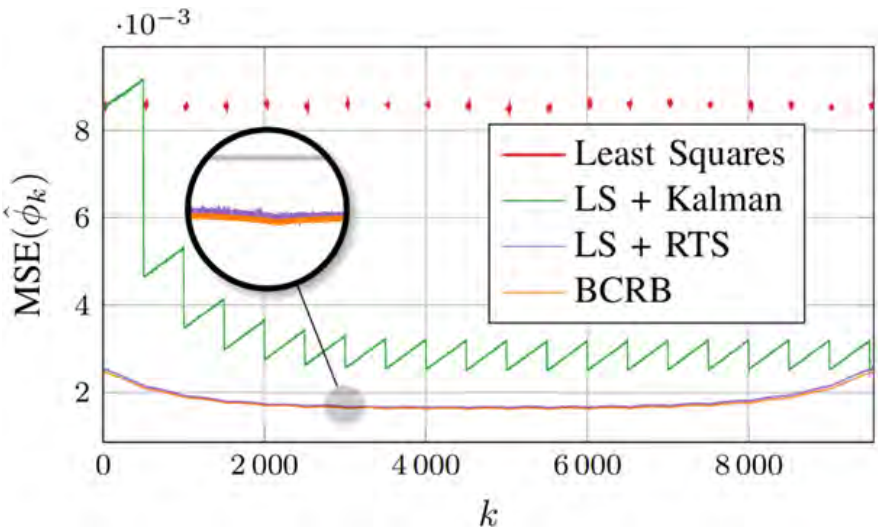


Figure 5-24: MSE of the estimators and BCRB over the sample index k within the observation frame, taken from [ZGD+22].

Table 5-14 Simulation parameters.

Transmit and Receive filter	Phase Noise Process	Transmit Signal
RRC	$\sigma_\xi^2 = 4K_2T\pi^2$	QPSK Symbols
Rolloff factor $\boldsymbol{\alpha} = \mathbf{0.5}$	$K_2 = 158$	Block size $N = 50$
Bandwidth $\boldsymbol{W_{Tx}} = \boldsymbol{W_{Rx}} = \frac{1+\alpha}{2T} = \mathbf{3.5\ GHz}$	$\sigma_{\phi_0}^2 = K_0/T$	Data block per pilot block $D = 9$
	$K_0 = -130\text{dB}$	

5.1.4 Potential of spatial combining and carrier aggregation in 6G waveform design

Using large array also gives possibilities to use the antenna, i.e., space domain to modulate the signal and design waveform. In such system, the waveform samples are generated by combining multiple signals over-the-air transmitted by different antennas. This basically mean that the overall waveform is designed in several components transmitted by different antennas or subarrays. The components can be generally derived to have different amplitude, different phase, or difference frequency. Other option is to have the components active in different time. Here the focus is selected on two potential concepts that are relevant in 6G context that are:

1. Spatially distributed hybrid beamforming enabled out-phasing transmitter
2. Utilizing carrier aggregation (CA) for link adaptation

In the following two sections, these potential concepts are explained and their applicability to 6G systems are discussed. Aspect (2) is discussed from the perspective of bandwidth vs link distance to study carrier-aggregation potential for link adaptation, while using band-selective subarrays are one way to realize the CA.

5.1.4.1 Hybrid-beamforming enabled out-phasing transmitter

In waveform design, the aim to reduce the PAPR is in the digital domain to enable low PAPR signal to feed the RF components, especially the PAs. Whereas this is a good target as such, this often leads to reduced spectral efficiency. From the RF perspective, it is important that signals feeding individual RF components have low PAPR. However, it should be noted that multiple low-PAPR signals combined can make one large-PAPR signal. In other words, by dividing the signal into multiple low-PAPR signals that feeds individual PAs and combining them then together can enable to make large PAPR signal with high efficiency from PA perspective. In the literature, this is often referred as so-called out-phasing transmitter [Chi35] or linear amplification with nonlinear components (LINC) [Cox74] where the signal

is distributed into two components and their output phase is then varied to achieve the varying envelope. Typically, the combining is performed using different types of conducted combiners in the circuit level to feed the varying envelope signal to the antenna. Such combining shown in Figure 5-25. In principle, any signal can be divided into two out-phase components that have constant amplitude. Let us simply assume $x(t)$ to be the input signal of the system that has varying envelope. Now, $x(t)$ can be written as a sum of two components $x_1 = \exp(1j\varphi_1(t))$ and $x_2 = \exp(1j\varphi_2(t))$ such as $x(t) = x_{phs}(x_1(t) + x_2(t))$. Note that $|x_1(t)| = |x_2(t)| = |x_{phs}(t)| = 1$. In out-phasing transmitter, the phases of the signal components are selected to be $\varphi_1(t)) = -\varphi_2(t))$. Now, by varying the phase, any value in the real axis can be produced. This makes it possible to generate varying amplitude by two constant envelope signals. The output phase of the signal is achieved by rotating the overall signal then also with a phasor $x_{phs}(t) = \exp(j\varphi_{phs})$. Hence, the overall signal consisting of amplitude and phase components $x(t) = |x(t)|\exp(j\arg(x(t))$ can be represented such that $|x(t)| = |x_1(t) + x_2(t)|$ and $\arg(x) = \varphi_{phs}$. This basically enables to have only phase-varying signal feeding each PA, while the amplitude is generated as a sum of the components at the output. This enables to have low PAPR per PA. In general, any type of waveform can be reproduced by using this technique.

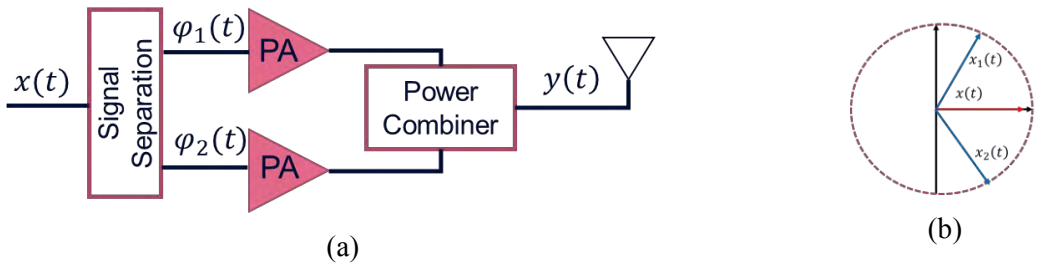


Figure 5-25: Illustration of typical out-phasing concept.

Using the concept above in very high centre frequencies such as upper mmWave region is challenging. In the circuit side, one of the challenges is the output-combiner that is typically rather lossy and hence reduces the overall efficiency and reduces the maximum achievable output power. One way to solve the problem is to replace the output combiner by simply individual antennas as proposed in [LR11]. This generally makes the output signal by combining two signals over-the-air. The interesting thing in such concepts is that it eventually makes the varying amplitude by (digital) beamforming and nicely combining waveform and modulation design, beamforming and RF design together. In large subarray-based hybrid beamforming systems this concept can be utilized by combining multiple subarrays in similar way. Such a concept that is shown in Figure 5-26.

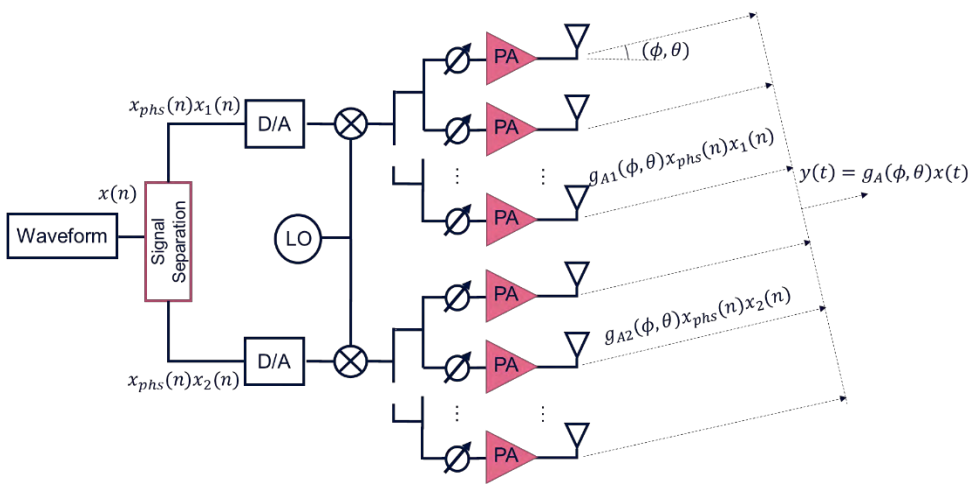


Figure 5-26: Illustration of over-the-air combined outphasing concept.

The example system consists of two subarrays having dedicated RF beamformers (phased arrays) that both have a dedicated digital signal chain, and mixer. The division to individual signal components can

be done in the digital side, while the over-the-air combining is then over two subarrays. In principle, this is basically generating the overall amplitude-varying signal by hybrid beamforming over two subarrays. Similar concepts are presented in the literature, for example, in [LXA+21] and [ASS+20], where in the latter one the concept is called as "outspacing" transmitter.

The basic requirement for applying the concept in practice is to have a multi-antenna transmitter with at least two digital signal chains. Applying the concept in 6G transmitter is possible as the architecture has multiple subarrays. Hence, it can be implemented as an alternative method to generate the higher PAPR signals. The main thing is that it does not require any extra RF hardware components. Added complexity comes from the signal separation to amplitude and phase components and amplitude separation to two out-phased components. This could be addressed in the waveform design side to generate the waveform to directly provide these two components. For example, in typical single-carrier modulations such as QAM this is straight forward [LR11]]) and performed in such a way that the typical pulse shaping is then performed after the signal separation for both out-phasing components. However, with typical multi-carrier transmitters, this is more challenging to address directly inside the waveform generation. For generic 6G transmitter architecture, out-phasing by hybrid beamforming can be enabled as a separate mode to use higher spectral efficiency modulation in some cases, while in other cases, the individual subarrays can operate in more typical modes serving users at dedicated directions with their own waveform or serving a single direction with one waveform same fed to both subarrays.

The main challenges of the technique are that it requires accurate calibration of the subarrays to make the arrays exactly phase coherent. This is expected to be challenging with very wideband waveforms. In addition, the other challenge is related to the fact that the quality (linearity) of the combined signal is highly direction dependent that is also analysed in the original reference [LR11] by showing that the achieved bit error rate is very sensitive for steering angle. The impact is similar than in directive linearization schemes presented in [Ter22, TAT+17]. This further increases the calibration requirements also to beam domain. On the other hand, direction dependent behaviour of the waveform can be also beneficial as discussed in [Ter22] to, for example, make it even more challenging to detect the signal outside the steering angle.

Note that the used technique does not, as such, increase the achieved P_{sat} in the steering angle as the maximum amplitude is achieved when the out-phasing is selected to be zero. The overall amplitude variations are, hence, achieved by using the array gain and the maximum array gain is achieved for the peaks of the signal. Note that same maximum P_{sat} for the signal peaks would limit any type of waveform eventually, but it is likely that in out-phasing mode it could be possible to go rather close to P_{sat} as the signal is per PA is expected to have low PAPR. Also, it enables to operate the PAs in more efficient mode that, for example, could help the thermal management of the amplifiers. Moreover, the PAs can be biased in more nonlinear mode for this case that can further improve the efficiency and hence reduce the power consumption of the power amplifiers.

5.1.4.2 Carrier aggregation for distance-aware link adaptation

CA is one powerful tool to allocate bandwidth for users, as well as to generate wideband waveforms as a sum of multiple more narrowband waveforms, which is referred as analogue multicarrier waveform [HEX21-D21]. When it comes to the very wide signal bandwidths envisioned for 6G, generating the wideband waveforms in smaller blocks makes it easier also to scale the power consumption of the data converters by switching sub-bands off for certain purposes. Numerous CA schemes for 5G systems has been already used and potentially some of them could also be proposed for 6G. Feasibility of these for sub-THz frequencies is yet not widely being studied. In practice, the aggregation can be performed in different domains at baseband, IF and RF. Even individual subarrays can be allocated for different component carriers (CC) and aggregated in spatial domain, but the price to be paid in that case is the reduced array gain. Such scenario could be the best fit for link conditions where the noise is not the limiting factor for the achievable data rate. When going to very wide signal bandwidths, one of the major challenges is that also the noise bandwidth will increase that naturally has a major impact on the link budget. This is essentially solved by higher number of antennas. Similarly, as discussed in Section

4.2 of this deliverable, also phase noise impact are dependent on the signal bandwidth in higher centre frequencies. Hence, bandwidth also play a crucial role in the achievable link ranges.

Carrier aggregation can be also effectively used to scale bandwidth for different link scenarios. The link-adaptation can be done basically in two domains: vary the modulation and coding scheme (MCS) or vary the bandwidth. Varying MCS is simply changing the spectral efficiency. However, highest spectral efficiency MCS is limited eventually at least by the EVM of the transmitter. The goal envisioned in [HEX21-D21] was to reach 100 Gbit/s with 100-200 m link distance. This can mean 15 – 40 GHz of signal bandwidth to reach with realistic MCS selection in line-of-sight conditions. To get the numbers into a context, the following simple analysis is conducted. Let us assume that the data rate with a certain SNR and bandwidth can be calculated as

$$DR = B[\log_2(1 + \gamma(B, d, f_c)) - SE_{loss}], \quad \textbf{(5-10)}$$

where γ is the effective SNR (caused by all RF nonidealities) in linear scale that depends on the signal bandwidth B, centre frequency f_c, and link distance d. The distance and frequency dependency are caused by path loss, while the bandwidth dependency is caused by the noise bandwidth of the receiver. The term SE_{loss} represent the loss in the spectral efficiency with respect to the theoretical Shannon capacity, which according to [Ter22: figure 2] is around 1.5-2bit/s/Hz for typical QAM modulations. By following the methodology presented in [TTP17, Ter22, LTJ19, LJT19], the contributions of the Tx EVM and Rx noise are combined to evaluate the effective signal to noise ratio achieved at different link distances, bandwidth, and centre frequency.

A simulation example is conducted to highlight how the bandwidth vs distance relation can be seen in different link scenarios. The simulation parameters are shown in Table 5-15. The TX power is taken as a single PA power following the Silicon germanium (SiGe) process saturated power estimated derived in Section 4.3 with roughly 9 dB output backoff from P_{sat}. Similarly, low-noise amplifier (LNA) noise figure is following the estimates of the SiGe process derived in [HEX21-D21]. Used antenna element gains are typical patch antenna gains and two different antenna configurations are selected. Note that in practice the number of antennas is limited by power consumption, area as well as Tx power limitations set by different authorities. The analysis is conducted for 150 GHz and 300 GHz centre frequencies with free space path loss assumption. The results for four different scenarios are shown in Figure 5-23. Each curve in the figure shows the fixed achieved data-rate with a certain link distance and bandwidth. The data rate is derived based on Equation (5-10). Let us first look the scenario at 150 GHz at Figure 5-23 (a) – (b). The clear observation when comparing cases 1 and 2 at 150 GHz is that if the power budget, mainly realized by higher number of antennas, satisfy well the SNR constraint of highest achievable spectral efficiency (limited by Tx EVM), there is no significant link-distance dependency on the data-rate in decent link distances up to 500 m. This is seen in case 2, where the curves for each target data-rate are flat lines. However, in case 1, the link distance dependency in large bandwidths becomes more evident. The scenario can be more clearly observed at cases 3-4 in 300 GHz shown in Figure 5-23 (c) – (d). When looking this as a function of bandwidth, there seems to be "best bandwidth selection" to achieve a certain data rate with a certain link distance. This dependency is mainly caused by the fact that the different link scenarios in cases 3-4 clearly are in the boundary of Tx EVM limited and Rx noise limited region. Hence, in these scenarios, varying both MCS and bandwidth, to serve a user at certain link distance would be possible. To make it clear, highest MCS should be generally always chosen if the SNR budget can afford it, but as the SNR depends also on the bandwidth, they should be varied both for best compromise in the link adaptation. In practice, the best way the allocate the bandwidth would be to vary the number of active CCs.

Table 5-15: Simulation parameters for data-rate analysis with different link distances and signal bandwidth.

Case	f_c	Tx EVM	Ptx	Ntx	Gatx	NF	Garx	Nrx
1	150	5	10	256	5	6	4	16
2	150	5	10	1024	5	6	4	64
3	300	5	-4.5	256	5	17	4	64
4	300	5	-4.5	1024	5	17	4	256

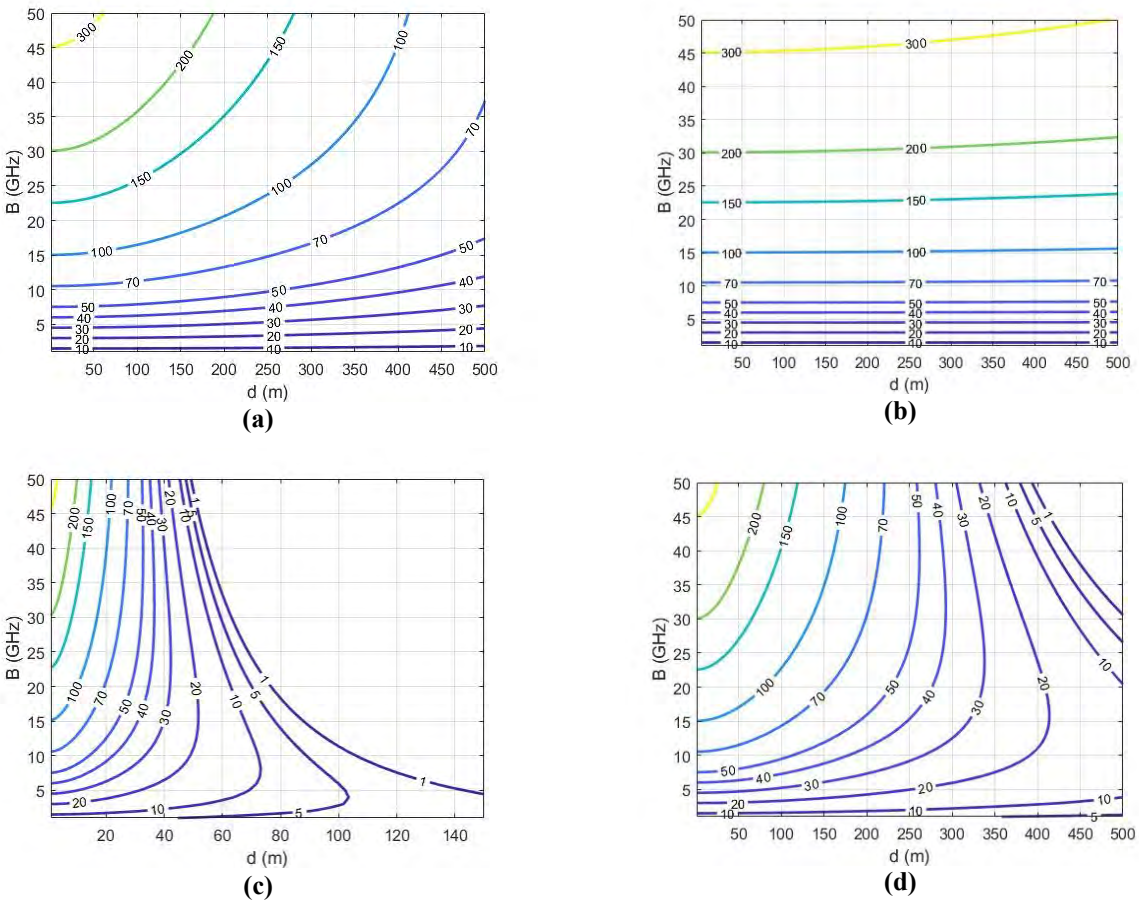


Figure 5-27: Simulated data rates at different bandwidth and link distance in four scenarios (a)-(d) derived in Table 14.

5.2 Beam management techniques

Beam management (BM) is defined as the processes of establishing as well as maintaining a wireless connection for a system utilizing any form of beamforming that requires beam alignment via testing different beams at least at one of the two end points. The general process is described by the finite state machine shown in Figure 5-28. It is obvious that each of these states as well as the state transition can be designed in different ways dependent on the system requirement. In the following discussion some aspects of the states and state transitions are introduced. They are also connected to the in academia well established procedure of beam training. Note that BM contains multiple aspects on top of beam training. This state machine has to exist per device pair, for simplicity the following description is assuming that a single device pair is present. The system states are always using a bold font. It is also important to notice that this finite state machine is existing at each of the participating devices. Sometimes one of the devices can transition earlier to another state. This can be resolved by other informing the other device of the state transition or by a later detection of the system state transition of the other device.

All the time there is no active connection of the system in the frequency range from 100 to 300 GHz the system is in the **BM idle** state. As soon as it reached an agreement on initial beam-alignment parameters with another device the system transitions to the state of **initial beam-alignment**. Note that this agreement can be explicit or implicit. Explicit agreement of these parameter would involve communication over another interface to agree on these parameters. Implicit agreement could involve an access point advertising it's service or any other type of detecting a potential communication partner. For example, visual detection via a camera is possible.

The tasks performed in the state **initial beam-alignment** can consist of a number of procedures. Combined they have the target that after the process is finalized communication via the interface utilizing the beamforming is possible. In a standalone 3GPP system this process would consist of first the initial beam-sweeping using the synchronization blocks with the UE measuring and selecting the optimal beam. The mobile device would afterwards use the knowledge gained from receiving the broadcast information to perform the random-access procedure. Afterwards a beam that can be utilized to establish a communication between these points is available, thus the mobile device and the access point have their **beams connected**. Note that it is also possible that the beam alignment fails, and the system would transition back to the **BM idle** state.

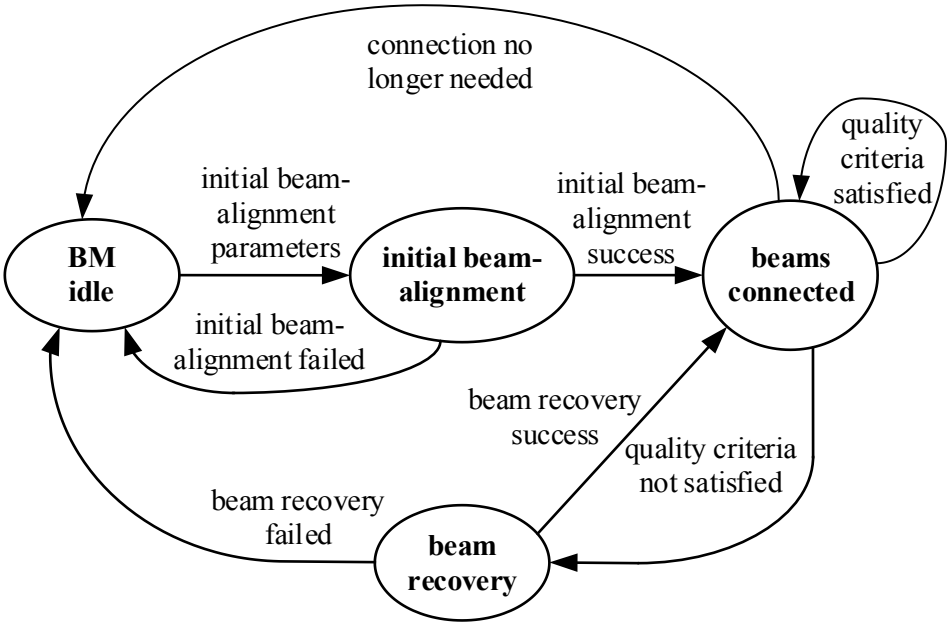


Figure 5-28: Beam management finite state machine.

By the time two devices have their **beams connected** to each other the link quality is monitored at both endpoints. If at any of the endpoints it is detected that the link quality is rapidly deteriorating or that the link is even failing it would transition to **beam recovery**. Note that it is possible that this is first or only detected by one of the participating devices. In this case the device that detects this state transition is informing the other one that **beam recovery** need to be started. In case that the connection is no needed the system can also transition to **the BM idle** state from the **connected** state. Inside the **beam recovery** state both the mobile device and the access point attempt to find a beam that fulfils the performance criteria. Based on prior information exchanges or even historic data different side information to assist the beam alignment are available relative to the procedure in the **initial beam-alignment** state of the system. Combining this observation of the shown finite state machine it is obvious that the main parts relating to the configuration and training of analogue beamforming are the initial beam-alignment, the beam recovery as well as the channel quality assessment during the beams connected state.

The beam training required for the initial beam-alignment is one of the challenges of the system design for communication systems in the frequency range of 100 to 300 GHz. This is because the antenna setup to achieve suitable coverage at this frequency can lead to a very narrow beamwidth. This in turns means that a lot of training is necessary. To give perspective of the required time, the following numbers illustrate an example. Assuming that the antenna panel at the access points covers a 120° in azimuth angle and 60° in elevation (considering a potential indoor scenario with a roof mounted antenna), there is a total of M = 1800 transmit beams if the beams have a width of 2°. Note that current 3GPP NR systems do not distinguish the beams per antenna panel but would need to train each beam at each panel separately, with this assumption the number of beams can even get significantly larger. Suppose that the antenna panel of the UE does cover 180° in azimuth and 90° in elevation (angled antenna around a corner of the device) with a beamwidth of 5° this results in N = 648. This does results in N*M = 1166400

beam pairs. Using the synchronization sequence with the broadcast channel as in NR for this beam training this would result in 4.56 s of just transmitting reference signals to train the beams, even if the subcarrier spacing is scaled o 3.84 MHz (2^8 * 15 kHz).

It is obvious that it will not be feasible to spend such a large amount of transmit power and time to align the beams of a single mobile device. Note that how severe this problem is going to be depends on the number of beams utilized by the access point as well as the mobile device. The number of beams depend on the coverage requirement as well as the channel conditions. At this point in time these plus the hardware capabilities of the future are not clear, thus accurate prediction and design of the related procedure need to wait till more information is available. The same observations made for the initial beam-alignment can be made for beam recovery. In this case due to the additional information of a past beam that was fulfilling the quality criteria is available in addition to the information available for initial beam alignment.

Different techniques to reduce the number of beams to be tested for initial beam-alignment as wells as beam recovery need to be defined. All these can be categorized into utilizing any time of side information available to the system. This side information can be historic observations, optical observations data, measurements at another frequency, positioning information, and information about the radio environment. Any combination of these is possible, however what is available for future systems operating in the frequency range from 100 to 300 GHz is at this point unclear. But as the system would be very inefficient with utilizing these, treating them as an integral part of the system design is paramount. For the quality monitoring of the connection with aligned beams a reactive or proactive approach is possible. In the reactive approach different quality criteria of the connections utilizing these beams are constantly monitored. In the case that the quality falls below a (potentially dynamic) threshold the network or the mobile device declares that communication with the current beams does fail. In the proactive approach in addition to all aspects of the reactive approach the system tries to anticipate a future degradation of the communication quality utilizing the current beam pair. This prediction can be based on any side information available to the network.

Combining this discussion about beam management it is clear that future communication systems utilizing the frequency range from 100 to 300 GHz will only work efficiently if the exploitation of side information is included at an early stage of the system design. One example of such side information are localization and sensing information as described in [HEX23-D33].

5.3 D-MIMO schemes

5.3.1 Layer-1 Mobility and Non-coherent Joint Transmission

Recent studies have shown distributed MIMO (D-MIMO) outperform traditional small cell and cellular massive MIMO networks in several practical scenarios. It has been conducted a thorough investigation and introduced centralized/distributed transmit precoding and receive combining techniques; various radio unit (RU) selection strategies that are motivated by practical implementation of D-MIMO systems. In previous deliverable, coherent joint transmission (CJT) in D-MIMO, which uses spatial repetition transmission (SRT), has been investigated for centralized and interference-aware distributed precoding with UE centric RU selection/clustering, and trade-off between implementation and deployment complexities has been shown. In addition to coordinated densification and macro-diversity overcoming the path-loss and removing the blocking effect for uniform performance, D-MIMO also helps realizing robust access links that supports Layer-1 (L1) mobility on high frequency bands where the propagation environment is more challenging. Due to time-varying propagation environment, channel aging effect becomes visible during mobility and high carrier frequency operation in D-MIMO networks.

This deliverable provides analyses on the impact of serving cluster update and its periodicity for downlink precoding in D-MIMO network including channel aging effect due to UE L1 mobility. Moreover, performance of non-coherent joint transmission (NCJT) using SRT among multiple RUs has also been evaluated for distributed and centralized precoding schemes with channel aging.

This section presents the following works: (a) Centralized and distributed precoding, i.e., zero-forcing (ZF) in particular, are studied with L1 mobility, i.e., serving RU subset update, see Figure 5-29. Distributed ZF done in clusters/subsets are interference-aware, i.e., consider minimizing the power interfering to other clusters [HEX21-D22]. (b) NCJT using SRT among multiple RUs is explored in a D-MIMO setup. (c) The impact of serving subset and precoder update periodicities are examined for both CJT and NCJT. The precoder must be updated when serving RU subset has been updated. The RU subsets, on the other hand, may not require as frequent updates as the precoder because the key factor for RU-UE association is slow fading which is mainly affected by the distance to the RUs.

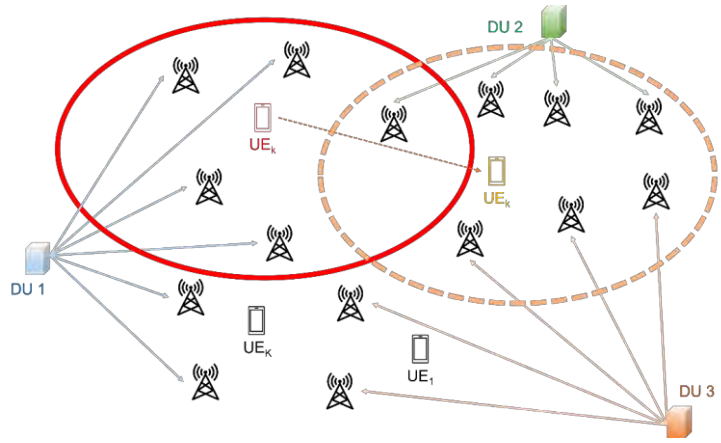


Figure 5-29: L1 mobility in D-MIMO with spatial repetition transmission.

Let L RUs coherently serve K randomly distributed and mobile UEs in a UE-centric way as shown in Figure 5-29, and received signal at k-th UE be represented in terms of channel gain ($\mathbf{H}$) and precoding ($\mathbf{W}$) matrices as

$$y = \sqrt{P}\mathbf{H}\mathbf{W}\boldsymbol{q} + \boldsymbol{n} = \sqrt{P}(\widehat{\mathbf{H}} + \widetilde{\mathbf{H}})\mathbf{W}\boldsymbol{q} + \boldsymbol{n} = \sqrt{P}\begin{bmatrix}\mathbf{H}_1^H & \mathbf{H}_2^H & \cdots & \mathbf{H}_K^H\end{bmatrix}\mathbf{W}\boldsymbol{q} + \boldsymbol{n}, \quad \textbf{(5-11)}$$

where $\boldsymbol{q}$ is the data symbol vector, $\widehat{\mathbf{H}}$ denotes the estimated channel matrix and $\widetilde{\mathbf{H}}$ is the bias matrix incorporating the channel aging impact, and $\mathbf{W}$ is calculated by using estimated channel matrix such that $\widehat{\mathbf{H}}\mathbf{W}=\mathbf{I}$ where $\mathbf{I}$ is the identity matrix; however, the received signal quality is still affected by channel aging. In case of centralized ZF (CZF), precoder for all UEs can be formulated as $\mathbf{W} = \widehat{\mathbf{H}}^H(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H)^{-1}$. In case of interference-aware distributed ZF (IADZF), UE-centric precoders per each subset can be found as $\mathbf{W}_k = \widehat{\mathbf{H}}_k(\widehat{\mathbf{H}}_k\widehat{\mathbf{H}}_k^H + \boldsymbol{\Lambda}_k)^{-1}$, where $\mathbf{H}_k = \begin{bmatrix}\boldsymbol{h}'_{k1} & \boldsymbol{h}'_{k2} & \cdots & \boldsymbol{h}'_{kL_k}\end{bmatrix}$ is the channel gain matrix for the k-th UE with L_k serving RUs for the k-th UE, and $\boldsymbol{\Lambda}_k$ is the interference covariance matrix [HEX21-D22]. The received signal is perturbed by a term of $\sqrt{P}\widetilde{\mathbf{H}}\mathbf{W}\boldsymbol{q}$ due to channel aging that is scaled by precoder, will eventually degrade the performance.

Performance Evaluation:

Performance of the D-MIMO system with L1 mobility and NCJT operating at 28 GHz is analysed in a DL indoor scenario, for different number of regularly deployed RUs, randomly distributed K=16 UEs and 1000 blockers. It is assumed that each RU and UE are equipped with single antenna, but it is straightforward to extend to the multi-antenna case. During L1 mobility, different serving subset update periodicities, i.e., 100 ms and 500 ms, are considered to update the serving RU subset. Other simulation parameters are given in Table 5-16. Perfect channel estimation and lossless, high-capacity FH is assumed. RF imperfections, hardware impairments, phase noise, power amplifier nonlinearities are omitted from the scope of this work.

Figure 5-30 shows spectral efficiency (SE) performance with different number of serving RUs with 100ms serving subset update periodicity. If more candidate RUs are deployed, better interference management can be done with increased degrees of freedom. Larger serving subsets can attain high performance for CZF in general and IADZF in case of dense deployments. Nonetheless, IADZF can provide less performance for a smaller number of deployed RUs since more interference starts to exist

in the UE specific clusters. As serving subset enlarges, i.e., number of serving RUs increases, performance difference between CZF and IADZF scales up.

Table 5-16: Simulation Model Specifications

Parameters	Model Specification
Bandwidth (MHz)	200
RU max transmission power (dBm)	13
Propagation Model	3GPP InH
Duplexing	TDD with 50% DL
Channel coherence time (ms)	100
UE speed (m/s)	5

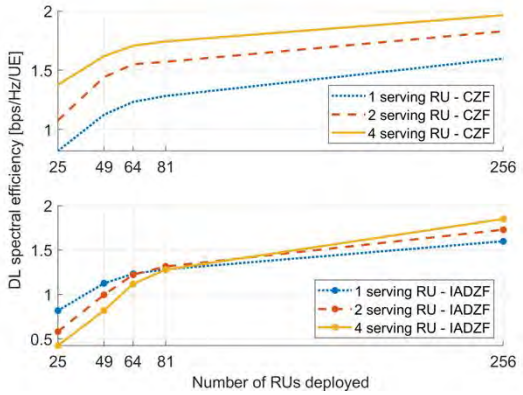


Figure 5-30: SE performance of CZF and IADZF with CJT for different number of serving RUs in a scenario with 16 RUs.

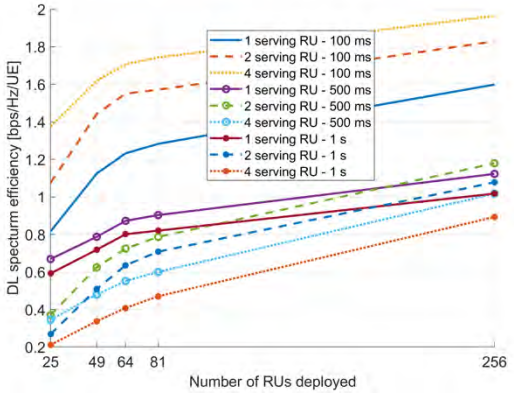


Figure 5-31: SE performance of CZF with CJT for different serving RU subset periodicities, e.g., 100 ms, 500 ms and 1 s.

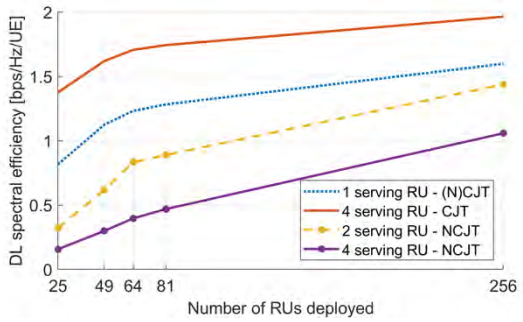


Figure 5-32: SE performance for CJT and NCJT schemes for different number of serving RUs in the scenario of 16 UEs when serving RU subset update periodicity is 100 ms.

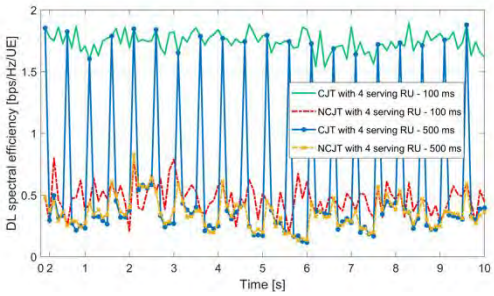


Figure 5-33: SE behaviour evaluated in 10s-time frame for CJT and NCJT schemes with serving subset size of 4 RUs and different subset update periodicities.

Figure 5-31 demonstrates the performance degradation if longer periodicity is adopted for serving subset update, where more performance loss has been observed for longer periodicities. And the degradation is higher in case of larger serving subsets. Frequent update is important to keep high performance for centralized precoding such that it loses much performance with slower updates. Figure 5-32 shows the impact of NCJT which attains less performance in multiple serving RU subsets relative to the CJT due to the not-optimized parameters. Provided that distributed precoding already performs worse, NCJT will cause relatively less performance loss compared with CZF. Figure 5-33 presents the performance w.r.t. time during mobility in case of serving subsets of 4 RUs in a region of 81 deployed RUs for both serving subset update periodicities. Herein CJT provides good performance only when precoder updates, whereas SE fluctuates with serious degradation under longer update periodicity that

is like NCJT performance. For NCJT, regardless of update periodicity, performance is substantially reduced. In short, frequent serving subset update is necessary for CJT, but not critical for NCJT.

5.3.2 Channel Estimation for D-MIMO in Dynamic Scenarios

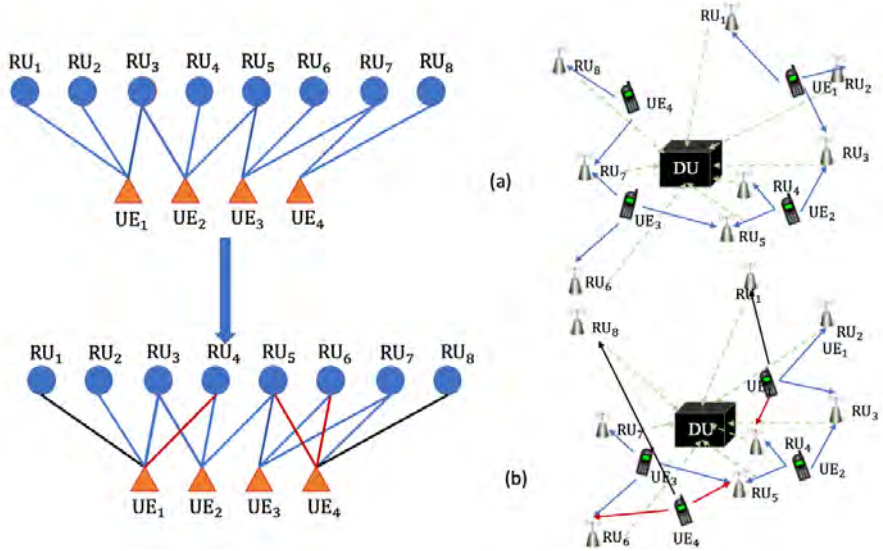


Figure 5-34: Modelling the topology of a dynamic D-MIMO network as a bipartite graph with partial connectivity.

To make the topology structure of dynamic D-MIMO more clearly, Figure 5-34 is an example of dynamic D-MIMO process, employing sparse bipartite graphs $\mathcal{G} = \{\mathcal{A}, \mathcal{U}, \mathcal{E}\}$, where $\mathcal{A} = \{a_1, a_2, a_3, \ldots, a_M\}$, $\mathcal{U} = \{u_1, u_2, u_3, \ldots, u_K\}$, denote the RU nodes and users nodes, respectively. In Figure 5-34 (a) and Figure 5-34 (b), the blue circle nodes represent RUs while the orange triangle nodes represent users. The upper two figures indicate the bipartite graph of the *t*-th snapshot, while the bottom one indicates the $(t + 1)$-th after structural modification. In comparison to the previous snapshot, in these four figures the red and black lines represent new and lost connections, respectively. Here, the edge set $\mathcal{E}$ is defined as $\mathcal{E} = \{e(a_m, u_k), m \in [M], k \in [K]\}$.

In this way, the received signal at *m*-th RU *t*-th snapshot during training phase can be written as

$$\boldsymbol{r}_m^t = \sqrt{\rho T_p} \sum_{k \in \mathcal{K}^m} g_{mk}^t \boldsymbol{s}_k + \sqrt{\rho T_p} \sum_{k' \notin \mathcal{K}^m} g_{mk'}^t \boldsymbol{s}_{k'} + \boldsymbol{w}_m, \quad \textbf{(5-12)}$$

where ρ is the transmit power of each user, and w_m is the i.i.d. additive AWGN at RU-*m* with $\mathcal{N}(0, \mathbf{I})$, g_{mk}^t denotes the channel coefficient between RU-*m* and user-*k* at *t*-th snapshot, $\mathcal{K}^m$ denotes the dominated users set connecting with the RU-*m,* and $\boldsymbol{s}_k$ is the pilot sequence of user-*k*.

In the aforementioned definition of the topology structure of D-MIMO, the topology graph modifications at the $(t + 1)$-th snapshot are discussed. This indicates that the connections between users and RUs have varied, and the assigned pilot sequence needs to be revised. The robust pilot assignment method indicates the specified pilot sequence strategy can be durable in a dynamic scenario; that is, even if the topology structure has changed, the pilot sequence could continue to operate on the updated graph if the distributed unit (DU) has received updated pathloss information.

MDS based channel estimation:

On the RU side, each RU receives a combination signal from at most T_p dominant pilot signals, allowing for the estimation of T_p dominant channels. Regarding the maximum distance separable (MDS) code generation process, it has been demonstrated that the Reed-Solomon code satisfies MDS code requirements. As one type of linear codes, Reed-Solomon and its derivatives are perhaps the most widely used codes due to their numerous advantages [Rot06]. The generator matrix is defined as

$C(K, T_p) = [c_1, c_2, \ldots, c_K]$, and each element in this matrix is an integer number that demonstrates the linear combination of pilots. In this way, the pilot sequence of the k-th user $\boldsymbol{s}_k$ can be specified as

$$\boldsymbol{s}_k = \sum_{j=1}^{T_p} \boldsymbol{\phi}_j c_{kj}, \tag{5-13}$$

where $\boldsymbol{\phi}_j$ is the orthogonal vector, and c_{kj} is the coefficient of j-th orthogonal vector. For example, in Figure 5-34, it is consided that all dominating channels (links in the bipartite graph) must be estimated, and the maximum dimension is same with the maximum number of edges connected to RU, which is 2 in the upper figure. A larger space than the initial snapshot is defined to enable the generated code to be available for more snapshot, e.g., $T_p = 3$. Then, generate the normalized MDS generated matrix as

$$\begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \boldsymbol{c}_3 \\ \boldsymbol{c}_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{7}{14} & \frac{1}{14} & \frac{6}{14} \end{bmatrix}. \tag{5-14}$$

Here, the 5th RU is demonstrated as the example. The received signal at the upper figure is specified as

$$r_5^u = \sqrt{\rho T_p} g_{52}^u \boldsymbol{c}_2 \boldsymbol{\Phi} + \sqrt{\rho T_p} g_{53}^u \boldsymbol{c}_3 \boldsymbol{\Phi} = \sqrt{\rho T_p} g_{52}^u \boldsymbol{c}_2 \boldsymbol{\phi}_2 + \sqrt{\rho T_p} g_{53}^u \boldsymbol{c}_3 \boldsymbol{\phi}_3, \tag{5-15}$$

and once the topology changes, the received signal at the bottom figure becomes

$$\begin{aligned} r_5^b &= \sqrt{\rho T_p} g_{52}^b \boldsymbol{c}_2 \boldsymbol{\Phi} + \sqrt{\rho T_p} g_{53}^b \boldsymbol{c}_3 \boldsymbol{\Phi} + \sqrt{\rho T_p} g_{54}^b \boldsymbol{c}_4 \\ &= \sqrt{\rho T_p} g_{52}^b \boldsymbol{c}_2 \boldsymbol{\phi}_2 + \sqrt{\rho T_p} g_{53}^b \boldsymbol{c}_3 \boldsymbol{\phi}_3 + \sqrt{\rho T_p} g_{54}^b \left(\frac{7}{14} \boldsymbol{\phi}_1 + \frac{1}{14} \boldsymbol{\phi}_2 + \frac{6}{14} \boldsymbol{\phi}_3 \right), \end{aligned} \tag{5-16}$$

As a result, both channels of upper and bottom figure at 5th RU can be resolved by using minimum mean square error (MMSE) or least square (LS) channel estimation methods. The MDS-based pilot assignment approach modifies the dynamic pilot assignment from dynamically assigning the pilot sequence to the selected solvable channel for each RU. In other words, the RUs receive sufficient signal to solve T_p dominating channels, which may be predicted using a designed estimator according to the graph structure e.g., MMSE. In fact, the criteria of the MDS-based pilot sequences are substantially less stringent than those of general orthogonal connection-based pilot assignment techniques, yet the MDS-based pilot sequence is remarkably robust. On the one hand, the dimension of the MDS code determines the highest capability of estimating the channel from the received signal, however, the property of the MDS has no bearing on the connection structure. Thus, a designed estimator may estimate the desired channels. Moreover, this non-orthogonal pilot sequence eliminates the restriction on the number of users in this cell. On the other hand, because each RU receives the non-orthogonal pilot sequence concurrently from all users throughout all timeslots, the interference from all other weak links cannot be disregarded. The detailed algorithm has been shown in here.

Algorithm 1: MDS-based Channel Estimation

1: Generate connection graph based the pathloss map.

2: According to the connection, generate the dominant users sets $\mathcal{U}_m^t$ for each RU

3: Define the pilot dimension T_p

4: Generate the RS code $\mathcal{C}(K, T_p)$, and sending pilot sequence to Rus

5: Output estimated channel of t-th snapshot

6: Once the connection changes, update the dominant users sets as $\mathcal{U}_m^{t+1}$

7: Channel estimation of $(t+1)$-th snapshot

Simulation result

An open square situation is considered, where $M = 200$ APs and $K = 100$ users with a single antenna are randomly distributed throughout a 1 km $\times$ 1 km square region. Modelling the large-scale fading coefficient β_{mk} as [BD16], the MDS-based method is compared with existing methods, such as semi-random [NAY+17], information-based k-means (IBKM) [CZE+21], and the sequential maximum weighted induced matching (sMWIM) [YYC22], respectively. Regarding the dynamic system concept, the topological structure is assumed to be dynamic. In this work, $N = 100$ snapshots are used in which the RU and user locations are randomly regenerated for each snapshot. Due to the sparse channel of the mmWave model, it is assumed that 10% connection in sMWIM unless otherwise specified. The pilot dimension T_p of the semi-random method sets to 6, and for the RU selection of structured policies, each RU maximizes the selection of 6 users. As there are more users in this cell, the value $\kappa = 3$ of sWMIW is adopted in this context. To generate the bipartite graph and determine the MDS-based method parameter, it is assumed that each user links $N_{in} = 8$ APs with stronger connections in the bipartite graph, and the pilot dimension for MDS is $T_p = 8$.

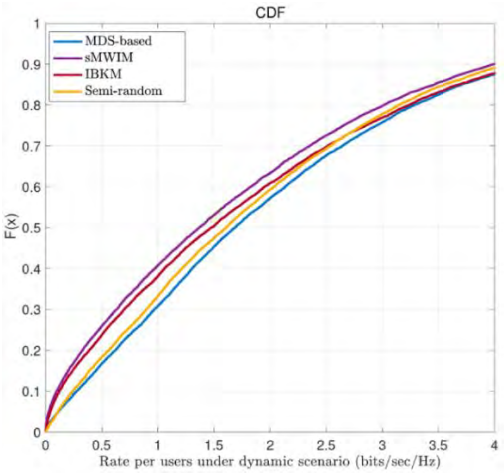


Figure 5-35: CDF of the downlink achievable rate per user under the dynamic scenario.

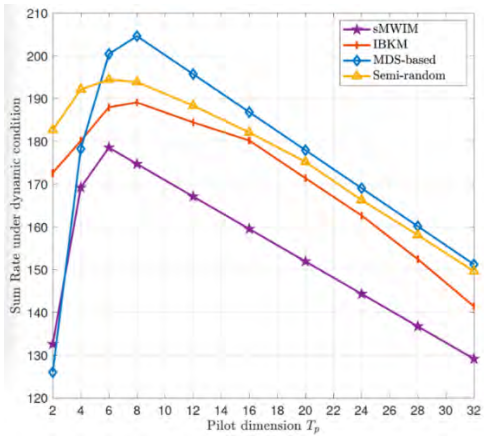


Figure 5-36: Rate versus pilot dimension under the dynamic scenario.

To evaluate the fairness of the baselines, it is assumed that the users utilize the same pilot sequence despite the changing topological structure, and that the DU modifies the pathloss information that is used for MMSE channel estimation. Figure 5-35 depicts the cumulative distribution function (CDF) of the downlink achievable rate per user of our MDS-based method in comparison to other benchmarks. In this figure, the performance of the proposed MDS-based approach exceeds that of the considered benchmarks, indicating that this method can sufficiently detect dynamic impacts. In contrast to the outcome of a static condition, semi-random provides superior robustness performance to the sWMIW and IBKM methods. This is owing to the fact that semi-random is not dependent on the system structure.

To illustrate the performance clearly, Figure 5-36 depicts the sum rate versus the pilot dimension T_p for all pilot assignment algorithms under dynamic conditions The MDS-based technique has the highest sum rate when the pilot dimension is $T_p = 8$, while the semi-random method has the best performance when the pilot dimension is small, i.e., $T_p = 2,4$. Unlike the performance under static conditions, the robustness of other techniques that depend on the connection structure presents a worse performance.

5.3.3 Analog D-MIMO: Conformance Testing and Validation

The relevant minimum RF characteristics and performance requirements for NR BSs are described in the 3GPP TS 38.104 [38.104]. Furthermore, for testing purposes, the 3GPP TS 38.141-1 [38.141] provides conducted test methods, relaxed requirements accounting for test tolerances, initial conditions, test procedures, among others. Appropriate test signals as defined in 3GPP 5G NR technical specification are used. 3GPP defines, in technical specifications (TSs), RF conformance test methods

and requirements for 5G NR BSs comprising the transmitter and receiver characteristics and performance testing. Thus, a technology must fulfil all mandatory RF tests and requirements defined by 3GPP to be considered NR compliant.

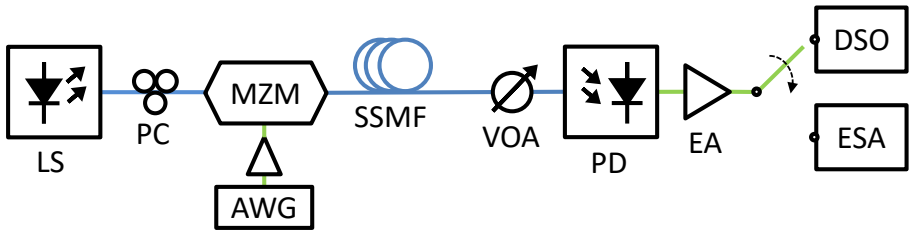


Figure 5-37: Experimental setup block diagram.

Figure 5-37 shows the block diagram of the experimental setup used for conformance testing, see [Puerta2022] for further details. An external cavity laser (ECL) at 1550.9 nm with a linewidth < 100 kHz and 16 dBm output power is used as light source (LS). An arbitrary waveform generator (AWG) operating with a sampling frequency of 50 GSa/s and 10 bits of vertical resolution generates the different NR test models (TMs) analogue signals. Common to all TMs, a NR carrier of 100 MHz (NRB = 273), SCS equal to 30 kHz, and carrier frequency of 3.55 GHz are used. The mobile C-band spectrum is chosen due to its large bandwidth available and its leading role in current 5G deployments [9]. The generated NR analogue signals modulate the LS output by means of a single- drive Mach-Zehnder modulator (MZM) with a Vπ equal to 4.2 V. The ARoF signals are generated biasing the MZM close to the quadrature point to get the best results. A polarization controller (PC) is used to adjust the light polarization at the MZM input. The output of the MZM is sent through 400 m of standard single-mode fibre (SSMF) which is enough for some of the envisioned implementations discussed in the introduction, e.g., D-MIMO. After fibre transmission, an optical variable attenuator (VOA) sets the received optical power and optical heterodyne detection is performed in a photodetector (PD) with an analogue bandwidth of 9 GHz and a responsivity of 0.8 A/W. After photodetection, the signal is amplified by an electrical amplifier (EA) with a gain of 22 dB and a noise figure (NF) of 6 dB. An electrical spectrum analyser (ESA) with an analogue bandwidth of 46 GHz is used to measure the ACLR of the ARoF signals using the parameters and procedure specified in [38.141]. To measure the EVM, the received signals are stored in a digital storage oscilloscope (DSO) operating with a sampling frequency of 10 GSa/s and 8 bits of vertical resolution for further offline digital signal processing (DSP). Using this setup, the feasibility of using ARoF links for future mobile fronthaul is assessed and it is validated experimentally that ARoF links are compliant with the 3GPP RF mandatory EVM and ACLR transmitter requirements. [GMM+22].

5.3.4 Comparison study of NCR and RIS

This work presents a comparison study of network-controlled repeater (NCR) and reconfigurable intelligent surfaces (RIS), and reveal some key features of both technical enablers.

System Model

To further illustrate the concept of the NCR, the network model presented in Figure 5-38 is considered, with an NCR-assisted multiple-input-single-output (MISO) DL system, where the base station (BS) is equipped with N_b transmit antennas transmitting to a single-antenna UE. To improve the system performance in the presence of, e.g., blockage, one NCR with N_n antennas on both the BS- and access-side of the NCR is deployed. The direct BS-UE link is ignored, due to, e.g., blockage. In this way, the power of the useful signal received at the UE can be expressed as

$$S = Pg\gamma_{BN}\gamma_{NU}, \quad \textbf{(5-17)}$$

where P is the transmit power at the BS and g is the amplification gain applied by the NCR. Also, in such a DL system,

$$\gamma_{i,j} = r_j H_{i,j} t_i, \quad \textbf{(5-18)}$$

represents the combined effective channel from node i (the BS or the NCR depending on the considered link) to node j (the NCR or the UE depending on the considered link) with transmit and receive beamforming, i.e., t_i and r_j, respectively. Also, $H_{i,j}$ denotes the channel between transmitter i = {b, n}, and receiver j = {n, u}. Assuming perfect channel state information (CSI), the transmit and receive beamformer can be obtained by, e.g., maximum-ratio transmission. The total noise power received by the UE can be calculated as

$$NP = \sigma_1^2 + \sigma_2^2 g \gamma_{NU}, \quad \textbf{(5-19)}$$

where σ_1^2 and σ_2^2 are the variance of the AWGN at the UE and NCR, respectively. In this way, the received signal-to-noise ratio (SNR) at the UE is *S/NP,* and the achievable rate is given by

$$R = B \log_2(1 + SNR), \quad \textbf{(5-20)}$$

with B being the channel bandwidth.

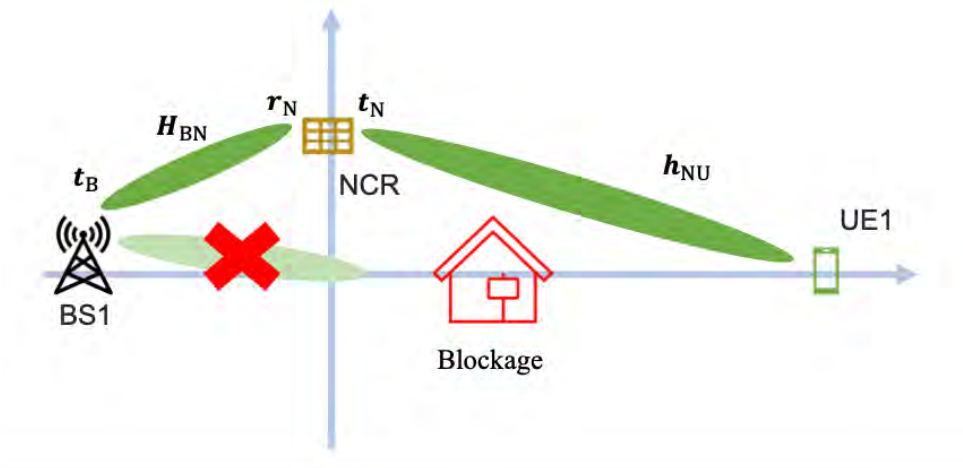


Figure 5-38: NCR-based DL system.

For RIS, a general model presented in, e.g., [ZDS+21], is adopted, and using alternative optimization to jointly optimize the precoder at the BS and the phase shifter at the RIS. With a cheap node, however, hardware imperfections may affect the reflection quality of the RISs. Particularly, it is likely that, in practice, RISs may provide an imperfect reflection because of different hardware impairments (HWIs), such as non-linear amplifier, phase error, quantization noise, where the performance of the RIS may be degraded. In this study, theHWI impairment models in [XWW+21] are used. Evaluation of different HWI models such as the ones in [GDS+22] is left as future work.

Results and Discussions

This section presents two highlighted results on the NCR and RIS. In both figures, the scenario consists of one BS with N_b = 16 antennas, one RIS with M = 100 elements/one NCR with 8 antennas at each side. Unless otherwise stated, the maximum output power constraint at the NCR is set to 40 dBm while the amplification gain is 90 or 100 dB. The carrier frequency f_c is set to 28 GHz with an ideal 1 GHz channel bandwidth Q. The noise power is set as -174 dBm/Hz with 10 dB noise figure. The antenna gains of the BS, RIS/NCR, and the users are set to 18 dBi, 18 dBi, and 0 dBi, respectively. For the HWI-affected rate [XWW+21], the following parameters are used; $\kappa_r = \kappa_t = 0.005$, $\alpha = 1$, and $\varphi_{BU} = \frac{\pi}{4}$.

Figure 5-38 gives an illustration of how the NCR helps the network in the presence of blockage assuming no additional BSs are helping to bypass the blockage (In practice, the neighbour BSs may also help in bypassing blockages [GMA+22]). As demonstrated in Figure 5-39, with the mmWave band, the blockage affects the UE achievable rate significantly. However, the presence of the NCR helps to bypass the blockage and avoid large performance drop. For instance, with the considered parameter settings of Figure 5-38, the presence of the NCR improves the UE achievable rate by 800%, compared to the cases with no NCR.

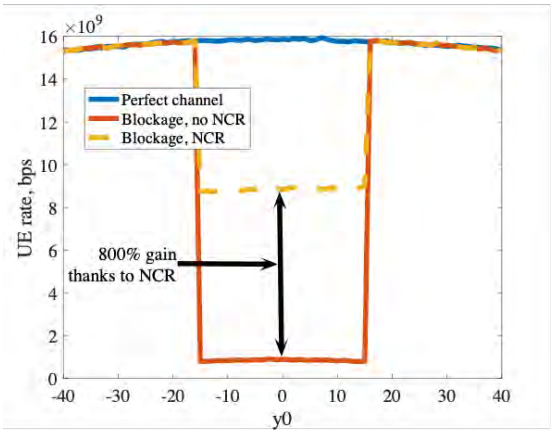


Figure 5-39: An illustration of how NCR helps the network in the presence of blockage.

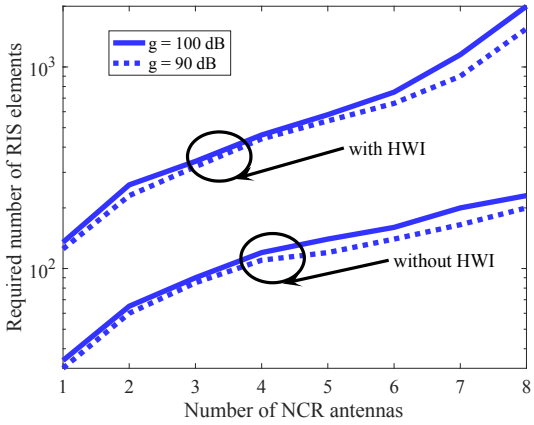


Figure 5-40: The required number of RIS elements to achieve the same performance as in the NCR-assisted network.

Figure 5-40 shows the required number of RIS elements to achieve the same throughput as in the NCR-assisted setup. Here, the results are presented for both cases with ideal and HWI-affected RIS models. Also, the transmit power at the BS and the maximum output power constraint at NCR are set to 43 dBm and 40 dBm, respectively with 90 dB/100 dB NCR amplification gain. As demonstrated, to achieve the same performance as in the NCR-assisted setup, a large number of RIS elements are required, and the required numbers increase with the NCR amplification gain. For instance, compared to the cases with 7 antennas at the NCR with amplification gain of 100 dB, around 200 RIS elements are required to obtain equivalent performance as in the NCR-assisted setup, if the RIS is ideal. With HWI, however, the required number of RIS elements increases significantly, for instance, to more than 1000 RIS elements in the considered example.

"In the study, the concept and structure of NCRs were introduced, along with conceptual simulations and a comparison of the performance of NCR-assisted networks with cases using RISs.

5.3.4 Mesh based IAB with multi-connectivity Access

5.3.4.1 Problem formulation

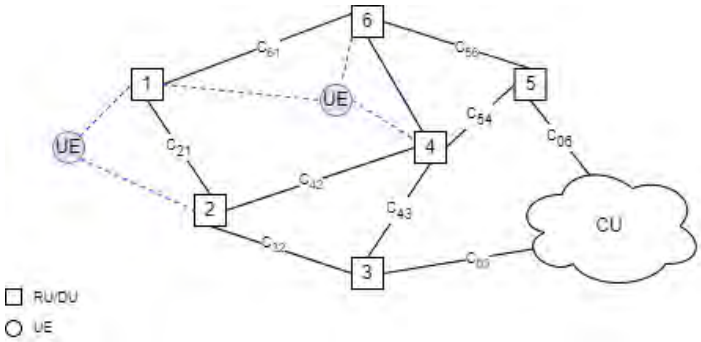


Figure 5-41: Generalized mesh front/backhaul problem formulation.

Higher data-rates drive the need for higher signal bandwidth and thus the use of higher carrier frequencies. Links at high carrier frequencies (upper mmWave) however result in lower link-budgets [HEX21-D21] and thus lower reliability on the wireless links. Blockage also can become a major issue at higher frequencies. Potential techniques to address these issues are:

- Reducing the range of links by for example moving RUs closer to the UEs (D-MIMO)
- Introducing link redundancy in the network by using for example mesh architectures.

	Shared Access-Backhaul Resource	Common Access-Backhaul Resource
Single UE-AP Connectivity	(1) Access links and back-haul links use separate frequency band(s), i.e., no interference between access and backhaul. A UE only connects to one RU at a time. The problem is reduced to an access scheduling problem and an independent mesh optimization as discussed in in [HEX21-D21]	(2) Access and back-haul links use the same frequency band(s). Access and backhaul need to be jointly optimized to avoid interference between the links. This is easiest done in a TDD based setup as presented in [HEX21-D22]
Multi UE-AP Connectivity	(3) Access links and back-haul links use separate frequency band(s), i.e., no interference between access and backhaul but a UE can connect to multiple RUs at a time. Link optimization in the access will affect mesh optimization.	(4) Access and back-haul links use the same frequency band and need to be jointly optimized. Connectivity of a UE to multiple RUs is supported.

Figure 5-42: Mesh resource optimization options.

In Figure 5-41, a generalized architecture, using D-MIMO in the access links (UE-RU/DU) and mesh in the back/fronthaul (RU/DURU/DU) is illustrated. The system problem discussed in this section is frequency and time resources can be used to optimize the network. Here, the problem is portioned into four approaches as summarized in Figure 5-42.

While option 1 (also referred to out-band IAB) is the simplest optimization problem, it does not make full use of the frequency resources (cannot share between access and backhaul). Option 2 (also referred to in-band IAB) tries to address this by operating on a single resource pool. This however results in a more complex scheduling problem especially in mobile scenarios [HEX21-D22]. Option three, which is the extension of option 1, but with the possibility of allowing multiple connection of a UE to the RUs is judged the most feasible approach to provide a reliable access and backhaul, while still being able to support changes in the network (mobility, weather, new nodes, etc.). Note that the complexity of this approach can be high if RUs care connected to different CUs. This option is discussed in the remainder of this section. Option 4 is only added for completeness and not discussed further.

The problem arising from option is two-fold:

1. Associate each UE with multiple RUs (two is chosen), to provide a reliable connection.
2. Given all UE to RU associations,
    a. optimize the routing of the user data to the core, and
    b. calculate fallback path through the mesh to the core, to guarantee robustness to single link failure events.

The dependent optimization of access and backhaul, classifies this approach as IAB, while the multi-connectivity of the UE, is a form of D-MIMO. The approach of creating fallback paths through the wireless network is not new. It is typically used in backhaul networks close to the core to provide error resilience. The problem formulation here is adopted to include the access link and to consider dense deployments, where potential backhaul links can interfere with each other.

5.3.4.2 Optimization and Results

The following system assumptions are made:

- Each UE is connected to one or two RUs. The RUs with best SNR are chosen. Latency from different routes can vary.
- A connection of a UE serving AP to the core is limited to bound the backhaul latency. This also reduces the routing complexity.

- The path cost of a UE is determined by the maximum cost of any segment in the path between UE and Core. The motivation for this choice is that the throughput of an end-to-end link between a UE and the core is limited by the link segment with the lowest throughput.

The considered network settings are summarized in Table 5-17.

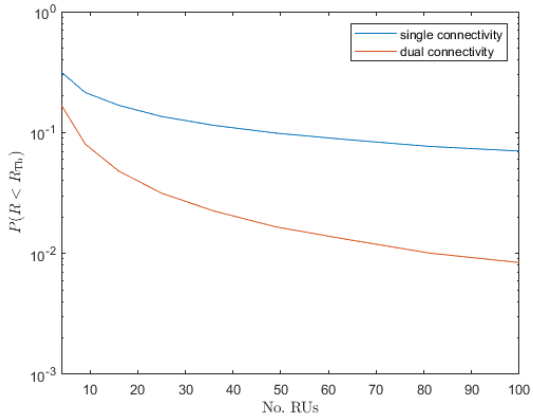


Table 5-17: Basic network parameters.

Carrier Frequency	60 GHz
Antenna Gain APs	15 dBi
Antenna Gain UEs	10dBi
Area	500m x 500m
UE distribution	Uniform
AP Distribution	Uniform grid with random offset
Core Connections	Uniform grid with random offset

Figure 5-43: Outage for Single and Dual connectivity in the UE-RU link. Outage is caused by shadowing/blocking which is a function of distance. Link-loss in mesh backhaul caused by interference. The number of APs with core connections was 9. Network simulation parameters are summarized in Table 5-17.

The graphs shown in Figure 5-41 show that for a low RU density, increasing the number of RUs results in an increased number of outage rate is improved since the distance between RUs an UEs is reduced. Allowing for two connections compared to only one from a UE, increases the robustness to shadowing and thus reduces the outage. The effect of adding more RUs to the network saturates for large number of RUs. Interference between the backhaul links starts to affect the reliability. Note that this simulation was performed for a fixed number of core-connections, even when number of APs was increased. Intuitively, adding core connections, is the same as reducing the number of mesh segments ('cell size').

In Figure 5-43, the number of hops required to reach the core from a UE serving RU is illustrated. For a given number of core connections, increasing the number of RUs does not reduce the number of hops further. For the chosen simulation setup moving with 4 Core connection RUs, increasing the number of RUs beyond 25, does not reduce latency (number of hops) anymore. Interestingly the distribution of hop-count starts to stabilize, which means that number of active links in the system starts to increase linearly with the number of RUs despite them operating in a mesh configuration. Adding more RUs mainly starts to improve reliability on the access links. Increasing the number of RUs with core-connections to 16, decreases the number of hops to reach the core network even for a higher number of RUs in the network. Compared to the same number of RUs and 4 core-connections, the number of hops reduces, and thus inter-RU interference in the mesh.

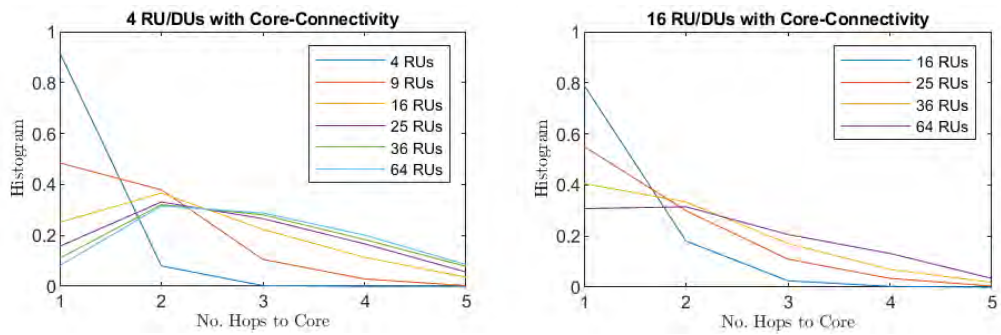


Figure 5-44: Distribution of number of hops for a different of core-DU/RUs and a different number of RUss. The algorithm used for calculating routes through the mesh was an iteratively applied Dijkstra optimizing for link SINR.

5.3.5 Constrained IAB Deployment Optimization

In this work, the effect of network planning on the service coverage of constrained IAB networks using different algorithms for constrained deployment optimization is studied. Here, the constraints are coming from either inter-IAB distance limitations or geographical/regulatory restrictions [MMH+22]. To further illustrate the implementation, consider the network model presented in Figure 5-45.

System Model

Consider the downlink communication of a two-hop in-band IAB setup where both access and backhaul links operate over the same mmWave spectrum. Here, the IAB donor and its child IAB nodes serve multiple UEs.

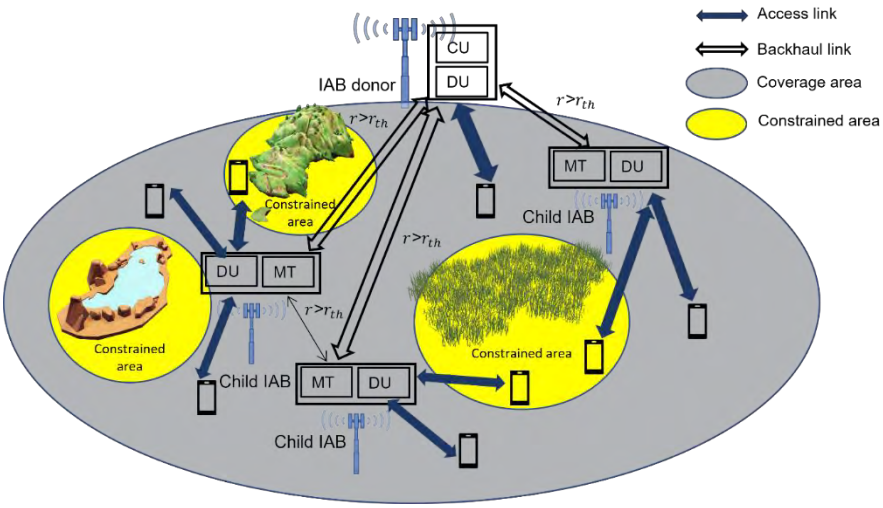


Figure 5-45: IAB network with deployment constraints.

Figure 5-45 shows an illustration of the IAB network with a minimum required distance between the IAB nodes and the IAB-MTs having gNB-like capabilities, with geographical constraints on node placement and the IAB. In this way, the power of the useful signal received can be expressed as

$$P_r = P_t h_{t,r} G_{t,r} L_{t,r} (||x_t - x_r||)^{-1}, \quad \textbf{(5-21)}$$

where P_t, $h_{t,r}$, $G_{t,r}$, $L_{t,r}$ denote the transmit power, small scale fading and pathloss according to 5GCM Uma close-in model respectively. The antenna gain model represents the sectored-pattern antenna array model described in [AMK+19], [MMC+20] where the main lobe antenna gain is significantly higher than the side lobe gain. In IAB networks the UEs can be served either by the child IAB nodes or IAB donor. In this context, following open access strategy and maximum received power rule the UEs are associated with the serving nodes. In a similar manner, the backhaul link association is determined based on the minimum pathloss rule. Moreover, the aggregated interference observed by the UE u due to the neighbouring interferers is given by

$$I_u = \sum_{j \in X_{i,u}\{w_u\}} P_i h_{i,u} G_{i,u} L_{i,u} (||x_i - x_u||)^{-1}, \quad \textbf{(5-22)}$$

where i represents the nodes excluding the associated node w_u of UE u. Similarly, the aggregated interference on the backhaul links can be calculated as given in [MMM+21]. The achievable UE rates are then calculated using Shannon's capacity considering sufficiently long codewords, which is an acceptable assumption in IAB networks [MMF+20]. Using the achievable UE data rates R_U, the study performs constrained deployment optimization such that service coverage given by

$$CP = \Pr(R_U \geq \rho), \tag{5-23}$$

is maximized where ρ is the defined threshold data rate. Here, the IAB placement with minimum inter-IAB distance requirement and IAB placement in the presence of constrained areas requirement are optimized using Algorithm 1 and Algorithm 2 respectively.

Algorithm 1 IAB placement with minimum inter-IAB distance requirement

With N_d IAB donors, N_c IAB child nodes inside the network area, do the followings:

I. Place the 1st node, $i = 1$, randomly in the considered network area.
II. Place the next node $i + 1$ where $i = 1, 2, 3, ..., (N_c + N_d - 1)$.
III. Find the minimum inter-node distances s_i between $(i + 1)$th node and each of other nodes.
IV. If any $s_i < r_{th}$, redistribute the last node $(i + 1)$th by repeating Steps II-IV until $s_i > r_{th}$.
V. For the obtained node locations, calculated the coverage. Then, proceed to Step I and continue the process for N_{it} iterations pre-considered by the network designer, saving the best set of node locations L_b among the considered solutions $L_j, \forall j, j = 1, 2, 3, ..., N_{it}$, which gives in the best value of the service coverage.

Return the set of the node locations in Step V as the optimal node location set.

Algorithm 2 IAB placement in the presence of constrained areas

With N_d IAB donors, N_c IAB child nodes and a set of constrained areas inside the network area, do the followings:

I. Place the IAB donors/IAB nodes randomly in the considered network area.
II. Identify the IAB node(s) falling inside the constrained areas.
III. For each of the nodes identified in Step II, redistribute the nodes.
IV. Proceed to Step II and continue the process until all IAB nodes fall outside the constrained areas. Save the set of locations as L_i. For each selected possible node locations L_i, compute the utility function CP_i, $i = 1, ..., N_{it}$. For instance, considering the service coverage, the objective function is given by (7).
V. Proceed to Step I and continue the process for N_{it} iterations pre-considered by the network designer, saving the best set of node locations L_b among the considered solutions $L_i, \forall i$, which gives in the best value of the utility function, e.g., service coverage.

Return the set of the node locations in V as the optimal node location set.

Here, greedy-based algorithms for optimizing the IAB placement with minimum inter-node distance constraint and geographical constrains are proposed respectively. Thereby, using a rejection-sampling method where multiple possible solutions are considered while satisfying the aforementioned constraints, the service coverage probability is maximized.

Performance Evaluation:

Performance of the optimized IAB system with deployment constraints at 28 GHz is analysed in a DL outdoor scenario, for different number of regularly deployed UEs, IAB child nodes and IAB donors. The general simulation parameters are given in Table 5-18.

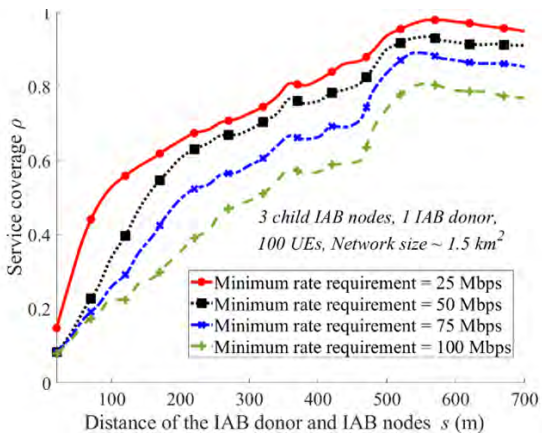


Figure 5-46: Service coverage as a function of the minimum distance constraint between the nodes for different UE data rate thresholds.

Table 5-18: Simulation parameters.

Parameters	Model Specification
Career Frequency	28 GHz
Bandwidth	1 GHz
Blocking	500 km-2
Noise power	5 dB

Figure 5-46 depicts a setup where the IAB donor is located at the centre and 3 child nodes are placed symmetrically, equidistant from the donor. As demonstrated in Figure 5-46, in such setup, the service coverage increases with the node separations, up to a point around 550 m. This is characterized by the effect of decreased interference between the nodes while having proper coverage in the area. Thereafter,

since the UEs can be located in between too far nodes, the coverage starts to slightly drop. In this way, there is an optimal distance between the nodes maximizing the coverage. Finally, the coverage decreases significantly with increased UEs minimum rate requirements due to the increased requirement of resources/nodes to serve such needs.

Figure 5-47 presents the results for the cases where the nodes locations are obtained only by considering the minimum distance between them or when the blockages and the backhaul links' qualities are also considered in the optimization. As depicted, the service coverage drops when the constraint becomes tighter, however, for all considered range of constraints, compared to hexagonal deployment, constrained deployment optimization increases the network coverage significantly. Indeed, prior knowledge of the blockages' locations helps to improve the network deployment optimization, especially when the UE density increases. Also, as observed, the effect of inter-node distance constraint on the coverage increases with the UE density. Here, the results are presented for the cases where the IAB donor has a main lobe antenna gain of 24 dBi and child IAB nodes have a main lobe antenna gain of 18 dBi while the side lobe gains are significantly lower, i.e., -2 dBi.

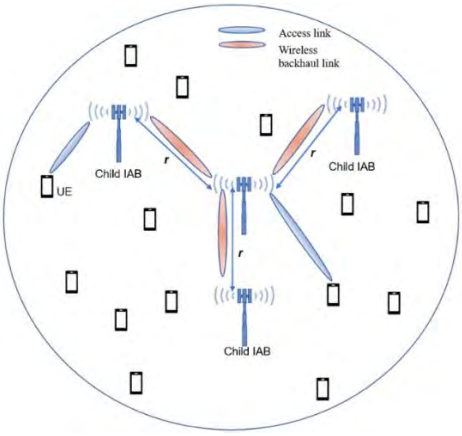


Figure 5-47: Symmetric IAB setup with donor at the centre.

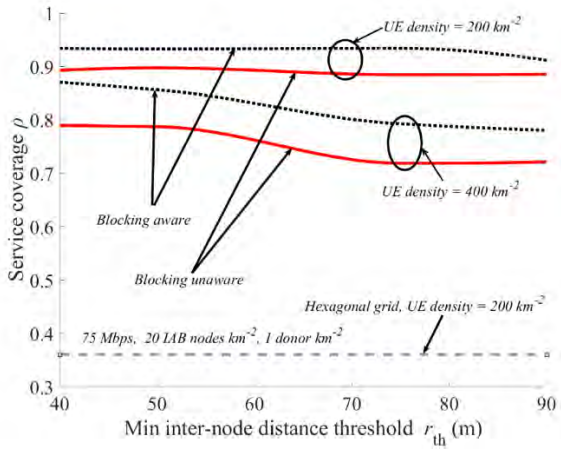


Figure 5-48: Service coverage as a function of the minimum distance constraint between the nodes in different UE densities for blocking aware and blocking unaware optimization scenarios.

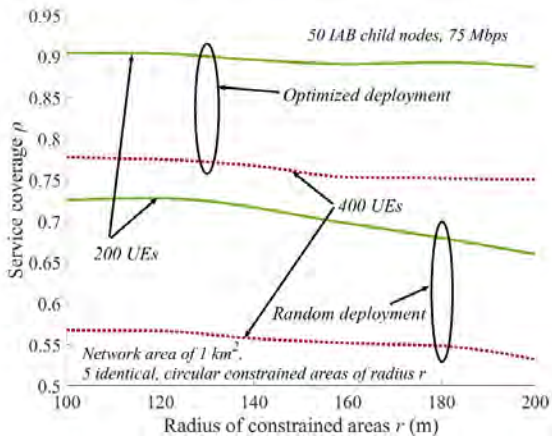


Figure 5-49: Service coverage as a function of the radius of the constrained areas for different deployment scenarios.

Figure 5-48 verifies the effect of geographical constraints, on the coverage of IAB networks. Particularly, the coverage of the deployment optimized IAB networks in the cases where, the IAB nodes are unable be placed in constrained areas either due to geographical or regulatory restrictions is studied. Here, the results are presented for a network consisting of five circular constrained areas of radius c. As demonstrated for low geographical constraints, network performance is not affected by the deployment constraints. However, with the presence of large area constraints, the performance, i.e., service coverage

decreases. This is intuitively because there is an increased chance of low coverage for users within the constrained areas when there are larger constraints for IAB placement. Also, since the IAB nodes get packed outside the constrained areas, interference levels for the UEs outside the constrained areas increases resulting in a further decrease in coverage. Finally, proper network planning boosts the coverage significantly compared to random deployment. Also, compared to the case with child IAB nodes distributed randomly in the unconstrained areas, the effect on the coverage is less severe when the network is optimized for geographical constraints.

6 Radio link and system performance analysis

This section presents selected results of the system performance analysis. In the first section, the impact of deployment scenarios on the power consumption is evaluated. In particular, the number of base stations for coverage, transceiver configuration, and, the analogue-to-digital converter (ADC) resolution. Sub-array-based transceiver architecture is considered, and the performance is evaluated in terms of throughput, and power consumption for different settings. The second section focuses on the influence of the propagation channel on the link quality based on channel measurements in several environments. First, two parametrized pathloss models are discussed, and the corresponding values are determined for each environment. Then, the number of independent beams is determined corresponding to the investigated indoor and outdoor areas.

6.1 Impact of deployment scenario on power consumption

A system simulation has been set up to investigate possible area coverage and achievable throughput with different base station configurations in an indoor scenario similar to the airport hall or shopping mall scenarios measured by Aalto University [HEX21-D22]. In a second step, for each of the base station configurations a coarse power consumption assessment is given, based on the power consumption models shown in [HEX21-D22, Table 3-4].

The scenario considered here is an area of 30 m x 60 m, with a base station (BS) located at the ceiling and pointing to the floor. Two scenarios are compared, the first is a 2-BS scenario with BS located at 10 m height, and the second is a 4-BS scenario with BS located at 5 m height. The height of the BS is chosen so that the coverage area per BS is coarsely 50% of the area in the 2-BS scenario, and 25% in the 4-BS scenario. This leads to a height of 5 m for the 4-BS case and 6 dB reduced transmit power compared to the 2-BS case when assuming the same reception (RX) power at the ground floor in boresight direction of the base stations.

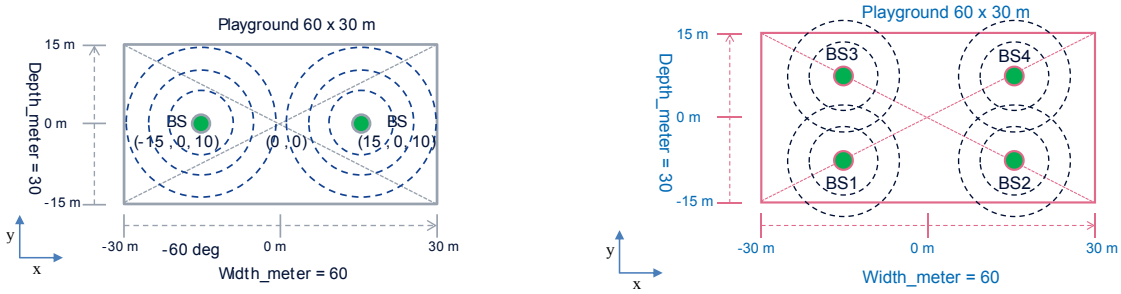


Figure 6-1: Deployment scenario with 2 BSs (left) and 4 (right) BSs illuminating the area from the ceiling.

The two scenarios are depicted in Figure 6-1. Each base station is equipped with a uniform planar array (UPA) with 8 x 8 antenna elements. Analog beam forming is assumed. To enable more than one radio frequency (RF) chains, the antenna can be divided into two or four subpanels with 8x4 and 4x4 elements, respectively. Each subpanel has one RF chain, so that 1, 2 or 4 RF chains are simultaneously active, and the respective number of simultaneous users can be served. Each subpanel in each configuration uses a grid-of-beams (GoB) with 9 x 9 beams covering an angle range of +/- 60° in both x and y dimensions. The subpanel configurations are shown in Figure 6-2.

With a system simulation the total throughput achievable within the deployment area, assuming 4, 8, 16 and 32 simultaneously served UEs, has been compared for the 2-BS and the 4-BS scenario (Figure 6-3). A total transmission (TX) power per BS of 41 dBm for the 2-BS scenario and 35 dBm for the 4-BS scenario has been applied. Each subpanel is assumed to serve 1 UE. For comparison also, the values for a fully digital GoB solution, which allows serving multiple UEs simultaneously, have been simulated. However, it should be noted that this is a theoretical configuration, since the number of RF chains for a full digital solution with large arrays might be prohibitively large. The purpose of this analysis is to investigate the power consumption for serving a specific indoor scenario. Hence, the

throughput in the total area is compared for the two configurations. Target is to find the most energy efficient deployment for a specific area with different number of simultaneous UEs. Therefore, the results are evaluated for the whole playground area and not "per cell", because also the number of cells is varying.

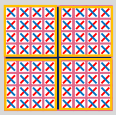
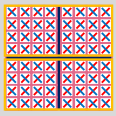
	GoB per subpanel (analog wideband beams)	Max. no. of simultaneously served UEs (signal streams)
A	 K=4 P=4 M_s=N/4 N=no. of elements	4
C	 K=2 P=2 M_s=N/2 N=no. of elements	2
D	 K=1 P=1 M_s=N N=no. of elements	1

Figure 6-2: Subpanel configurations with 1, 2 and 4 subpanels.

In Figure 6-3 results for the subpanel configuration variant D with one subpanel are shown. This means that for 2 BS only 2 user equipment (UEs) can be served simultaneously (light blue bar). For larger number of simultaneously active UEs the served UEs will change over time, but lead to similar average area throughput, while the throughput per individual UE goes down. For 4 BS up to 4 UEs can be served (dark blue bar). The throughput for 4 UEs is larger than with the 2 BS scenario, but also stays as number of simultaneously active UEs increases. In fact, per BS only one UE is scheduled at a time, due to the analogue GoB applied to the subpanels. In contrast, with a full digital GoB, where multiple UEs can be served simultaneously, the number of simultaneous UEs increases, and also the area throughput increases (orange and green bar). For reference, also the throughput with full digital array and eigen-beamforming is shown (grey bar, not fully comparable since there the UEs are assumed to have single antenna, in the GoB simulations the UEs have two antennas).

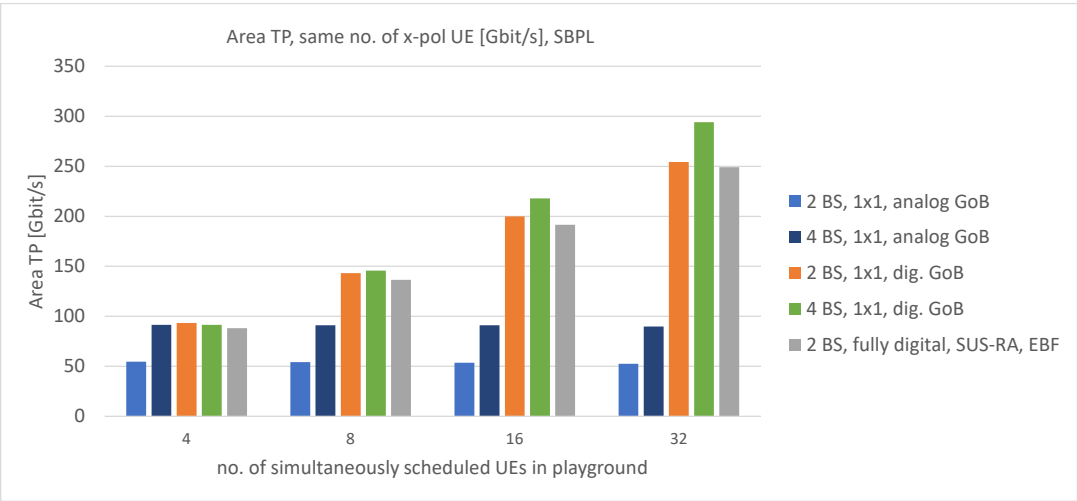


Figure 6-3: Area throughput for different number of simultaneously served UEs.

In Figure 6-4 the 4-BS scenario is shown, and the number of subpanels is varying. With 1 subpanel up to 4 UEs can be served simultaneously in the area (dark blue bar). With increasing number of UEs, the throughput is distributed among the higher number of UEs. When increasing the number of subpanels, an increasing number of UEs can be served, thus increasing the area throughput up to 8 UEs for 2

subpanels (light blue bars), and up to 16 UEs in case of 4 subpanels (magenta bars). The dark blue, green and grey bars correspond with bars of same colours in Figure 6-3.

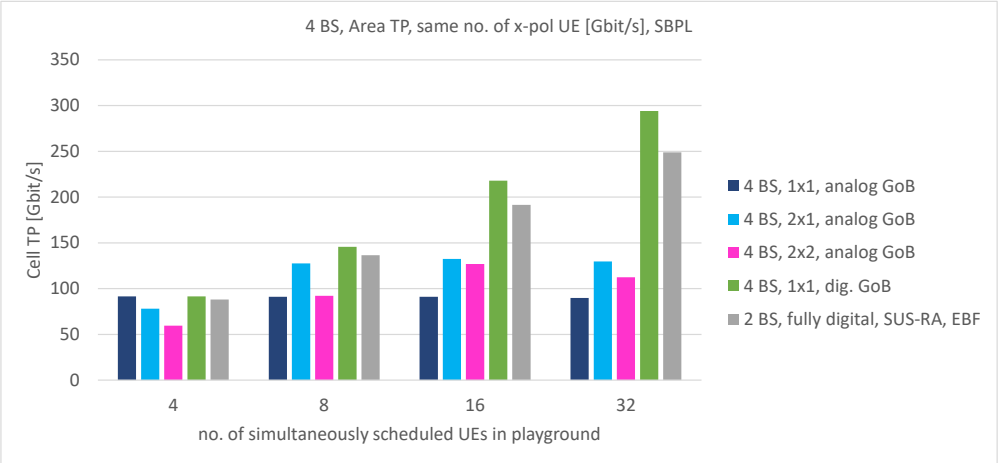


Figure 6-4: Area throughput for different number of simultaneously served UEs.

When looking at the UE throughput cumulative distribution function (CDF) (exemplary for 1 subpanel) over the deployment area shown in Figure 6-5, it can be seen that for 4 BS the area throughput CDF shows higher throughput than with 2 BS (red curves more to the right than blue curves).

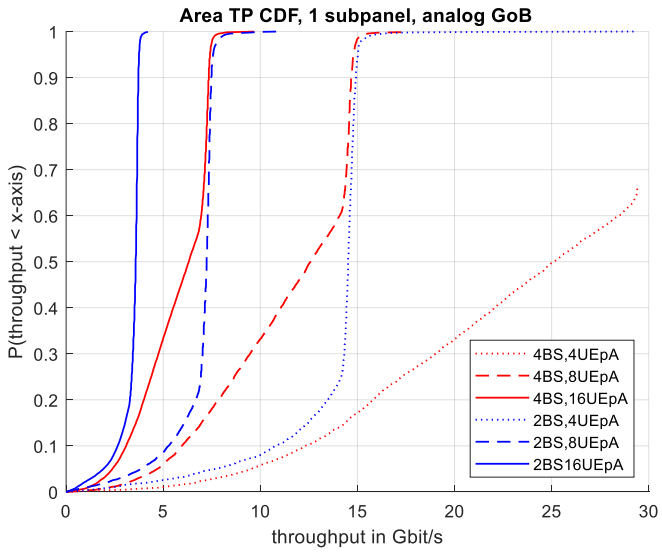


Figure 6-5: UE Throughput CDF for 2 and 4 BS over the whole deployment area, 1 subpanel and different number of simultaneously served UEs.

The next step is to assess the related power consumption of these configurations and bring it into relation to the throughput, to come to energy efficiency comparison. Basis for power consumption analysis is the power consumption model derived for antenna array hardware and high bandwidth data conversion as described in [HW21] and [HEX21-D22The power consumption of the RF frontend, including data conversion, is assessed. However, the power consumption model does not yet include baseband processing and is subject to further study. In Figure 6-6 the results for 2 and 4 BSs and also different ADC resolutions of 8 and 12 bits are shown.

It can be seen that for the 4 BS case, where the number of PAs is higher but the TX power can be lower, the overall power consumption is reduced although the area throughput has increased. Thus, a further increase of the number of BS could potentially further reduce the power consumption. But at the same time the number of RF chains increases, which leads to a trade-off and related optimum depending on the specific scenario.

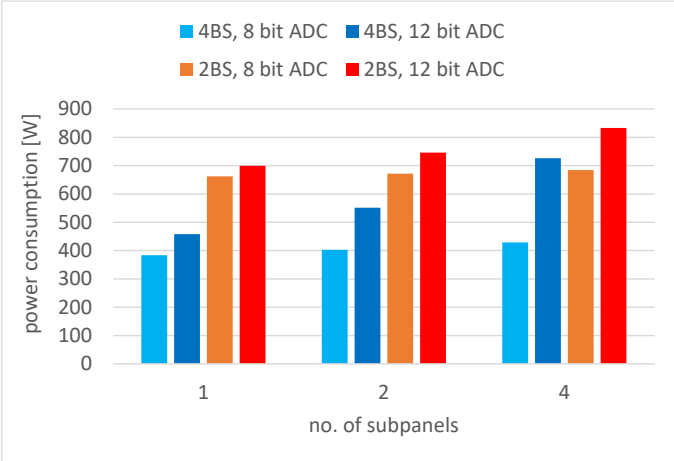


Figure 6-6: Exemplary power consumption for 2- and 4-BS scenarios and different ADC resolution.

6.2 Radio channel properties influencing link and system performance

This subsection describes an auxiliary part of the stored channel model detailed in Section 3. The stored channel model is complete without the following models of pathloss and number of independent beams.

6.2.1 Pathloss, angular and delay dispersions of outdoor and indoor radio links

The path loss, delay and angular spread models of a wireless channel are most influential characteristics of channel to radio links and systems. These parameters are estimated for the line-of-sight (LOS) and non-line-of-sight (NLOS) links in the entrance hall and three outdoor scenarios namely suburban, residential and city centre, described in [HEX21-D22, Section 7]. Two path loss models such as the close-in (CI) reference free space reference and the floating initial pathloss model, also called 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

$$PL^{\mathrm{CI}} = FSPL(f_c, 1\ \mathrm{m}) + 10n\log_{10}\left(\frac{d}{d_0}\right) + \chi_\sigma^{\mathrm{CI}}, \quad \textbf{(6-1)}$$

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

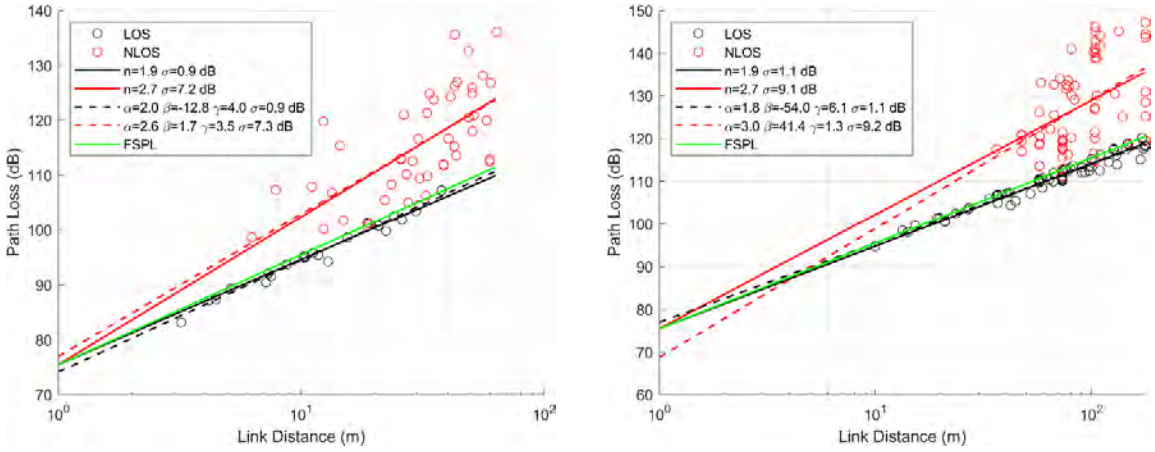
$$PL^{\mathrm{ABG}} = 10\alpha\log_{10}(d) + \beta + 10\gamma\log_{10}\left(\frac{f}{1\ GHz}\right) + \chi_\sigma^{\mathrm{ABG}} \quad \textbf{(6-2)}$$

where α is the distance-dependent loss coefficient, β is an offset coefficient, γ is the frequency-dependent loss coefficient, and $\chi_\sigma^{\mathrm{ABG}}$ is the shadow fading for the ABG model. The omnidirectional path loss for the indoor and outdoor scenarios, together with the fitted path loss models are plotted in Figure 6-7.

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 μ and standard deviation σ of angular and delay spread values are listed in Table 6-1: Angular and delay spread statistics in indoor and outdoor scenarios.. 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 [DKH21]. Similar to the path loss

modelling, the statistics of various spread values for the outdoor case include those from the three outdoor scenarios.



(a) indoor scenarios. **(b) outdoor scenarios.**

Figure 6-7: Path loss models.

Table 6-1: Angular and delay spread statistics in indoor and outdoor scenarios.

Scenario		AoD [deg]		ZoD [deg]		AoA [deg]		ZoA [deg]		Delay Spread [ns]	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
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 Center	**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

6.2.2 The number of independent beams of sub-THz link with a base station equipped with an antenna array

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 utilise 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. Hence it is interesting to study how many independent beams of practical beamwidth does the propagation channel support. Directional wideband propagation measurements mentioned in the section 0, i.e., (3-4), 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 [KGH+22]. The second method is based on measured single directional PADPs $P_q(\Omega, \tau)$ and a synthetic beam pattern $G(\Omega)$ defined in [3GPP17]. Only vertically polarized antennas were used in the channel measurement, therefore only single polarization is considered in the number of beams analysis.

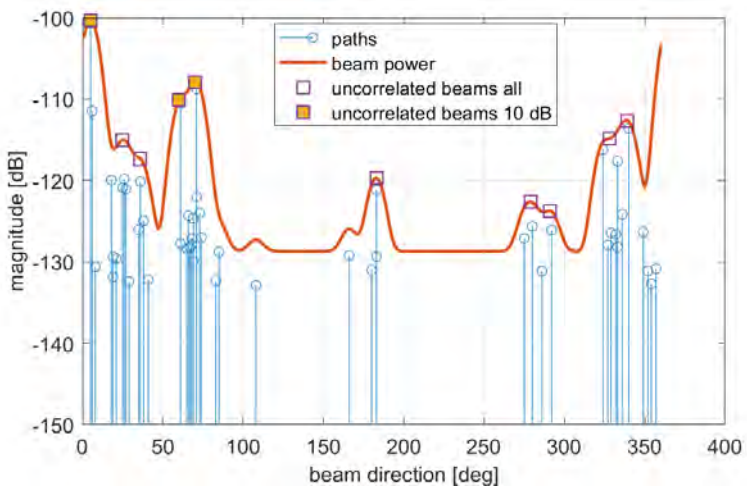


Figure 6-8 Measured path powers, the beam power and found independent beam azimuth directions of an example link.

Measured PADPs were evaluated using 10° half power beam width (HPBW), 10 dB dynamic range below the strongest beam, 2 GHz bandwidth (BW), and 0.5 correlation threshold between beams. An example PADP, beam power and identified beam directions are illustrated by blue circles, red curve and orange squares, respectively, in Figure 6-8. The identified independent number of beams in 132 measured indoor transmitter (Tx) and receiver (Rx) locations are shown in Figure 6-9. Figure 6-10 depicts empirical cumulative distribution functions (CDF) of the number of beams in 132 indoor and 157 outdoor locations using both 10, and 20 dB dynamic range. Median values of the number of beams are two in both environments using dynamic range of 10 dB, and using 20 dB they are five and three in indoor and outdoor environments, respectively.

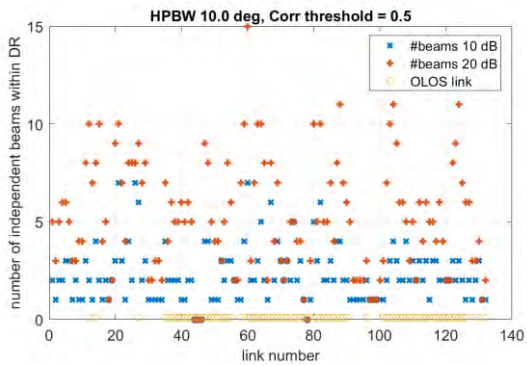


Figure 6-9 Number of independent beams in 132 indoor links.

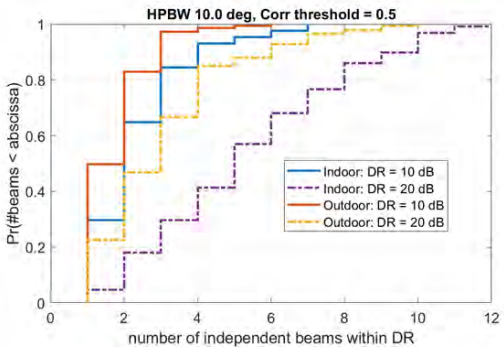


Figure 6-10 CDF of the number of independent beams in 132 indoor and 157 outdoor links using 10, and 20 dB dynamic range.

7 Conclusions

Sub-THz radio access network has the potential to realize use cases with extreme performance requirements in terms of data rate and latency. However, designing radio hardware at sub-THz upper Millimetre-wave frequencies (100-300 GHz) is challenging due to the complexity of multi-antenna implementations, hardware impairments, transmit power limitations, and propagation channel properties. A single radio configuration may not be suitable for all use cases and deployment scenarios, especially, when considering sustainability metrics. The radio design option depends on factors such as required data rate, communication range, and device type. Moreover, the physical layer modulation scheme and frame structure should fulfil the air interface latency requirements. Accordingly, the report proposes an approach for analysing radio systems, which involves considering multiple radio options based on different use cases. First, the requirements for each use case are determined, such as data rate, range, and number of users, and then used for deriving high-level technical requirements and communication scenarios. For each scenario, multiple radio options are proposed for different frequencies and device class.

The report discusses cellular-based and D-MIMO deployment options and highlights the potential of D-MIMO technology to optimize for coverage, enhance the capacity and reliability by cooperative transmission over multi-links. However, high data rates can increase processing costs for fronthaul and backhaul. The report explores various approaches to address this challenge.

A technical guideline for radio frequency (RF) transceiver design is presented with detailed discussions and evaluation example in each step. First, the technical requirements to achieve the data rate are presented in terms of bandwidth (BW), signal-to-noise ratio (SNR), and the challenges of analogue-to-digital convertors (ADC). Then, it discusses the requirements to achieve the SNR for a give range considering the RF nonidealities, and waveform impact on the power amplifier (PA) backoff. In addition to this analysis, it provides insight on the RF implementation aspects and strategies such as channelization and exploiting the concept of carrier aggregation (CA) to optimize the system. In each of this step, there can be several options to select from. Thus, other performance metrics, such as energy consumption and complexity, need to be considered to reduce the design space.

Based on the analysis, it is concluded that, for sub-THz frequencies, a hybrid architecture based on subarrays is a practical choice. The number of subarrays controlled with a full RF chain can be optimized depending on the number of simultaneous users and available spatial beams by the channel. Moreover, it enables the implementation of wide-band waveforms by aggregating multiple subarrays over the air. Furthermore, it provides control capability for power consumption by means of switching off the unneeded chains. However, insight analysis on beam management reveals that conventional beam tracking is inefficient due to the large number of beams to be probed with sub-THz narrow beams. Therefore, new approaches are needed to exploit available side information.

The report emphasizes the importance of hardware and channel modelling in the design process to consider non-idealities, compensate for path loss and exploit spatial beams. Several models for PA and phase noise are described and their accuracy is evaluated using measurement of lab components. A companion with state-of-the art models is introduced, proposing an extension of those models to cover wider frequency band above 100 GHz. In addition, the report highlights that channel measurements alone are insufficient to evaluate performance, as they depend on channel sounding and environmental factors. Therefore, data processing and the derivation of an extended channel model are critical for design. The modelling work in this report is complementary to the results presented in the second deliverable [HEX21-D2.2].

After designing the radio, the physical layer signal processing is evaluated considering one or more hardware models focusing on frame error rate (FRE). In the section of the waveforms, one approach is based on designing hardware-friendly modulation schemes, and another approach is based on developing digital techniques to estimate and mitigate the impact of hardware. For instance, OFDM variants such as DFTS-OFDM and SC-FDE are seen as potential candidates due to their low peak to average power ratio (PAPR) and resistance to phase noise, especially with optimized constellations

schemes. Moreover, these waveforms can also be easily integrated in the sub-array architecture based on spatial combining or carrier aggregation. On the contrary, zero-crossing modulation (ZXM) with 1-bit quantization, as an energy-efficient waveform, needs to cope with the phase noise by estimation and compensation.

A study in the report, focusing on achieving air interface latency, demonstrates that achieving 0.1 ms requires a short transmission time interval (TTI). However, the processing latency raises as a major challenge, especially, with extremely high data rates. This needs to be considered in the processing architecture design, considering I.O speeds, memory access speed, and power consumption.

Several performance examples demonstrate that optimizing deployment scenarios, transceiver configurations, and ADC resolution can significantly impact the power consumption and improve performance. Thus, the future work should focus on holistic optimization of radio design considering different degrees of freedom not only in the infrastructure but also in the signal processing, to ensure meeting the sustainability goals, while fulfilling the extreme requirements of emerging use cases.

References

[37.910] 3GPP TR 37.910 v16.1.0, "Study on Self Evaluation Towards IMT-2020 Submission (Release 16)"

[38.104] 3GPP, TS 38.104 V16.7.0: "NR Base Station (BS) radio transmission and reception," Apr. 2021.

[38.141] 3GPP, TS 38.141-1 V16.7.0: "NR, Base Station (BS) conformance testing, Part 1: Conducted conformance testing," Apr. 2021.

[38.211] 3GPP TS 838.211 V17.2.0: "NR, Physical channels and modulation (Release 17)"

[38.820] 3GPP, Technical Specification Group Radio Access Network; NR; 7-24 GHz frequency range (Release 16)," TR 38.820, V16.1.0, Mar., 2021

[3GPP17] TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," 3GPP, Tech. Rep. V14.1.1, Jul. 2017.

[AAK16] J. Ala-Laurinaho et al., "2-D Beam-Steerable Integrated Lens Antenna System for 5G E-Band Access and Backhaul," in IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 7, pp. 2244-2255, July 2016.

[AMK+19] S. M. Azimi-Abarghouyi, B. Makki, M. Nasiri-Kenari, and T. Svensson, "Stochastic geometry modeling and analysis of finite millimeter wave wireless networks," IEEE Trans. Veh. Technol., vol. 68, no. 2, pp. 1378–1393, Feb. 2019.

[ARA+14] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi and R. W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," in IEEE Transactions on Wireless Communications, vol. 13, no. 3, pp. 1499-1513, March 2014.

[ASS+20] B. G. M. van Ark, A. B. Smolders and P. F. M. Smulders, "Outspacing Phased Arrays for mm-wave 5G Base Stations," 2020 14th European Conference on Antennas and Propagation (EuCAP), 2020, pp. 1-5

[AZR92] I. Angelov, H. Zirath, and N. Rosman, "A new empirical nonlinear model for hemt and mesfet devices," IEEE Transactions on Microwave Theory and Techniques, vol. 40, no. 12, pp. 2258–2266, 1992.

[BD16] S. Buzzi and C. D'Andrea, "On clustered statistical MIMO millimeter wave channel simulation," arXiv preprint arXiv:1604.00648, 2016.

[BG03] S. Boumaiza and F. M. Ghannouchi, "Thermal memory effects modeling and compensation in rf power amplifiers and predistortion linearizers," IEEE Transactions on Microwave Theory and Techniques, vol. 51, no. 12, pp. 2427–2433, 2003.

[BP04] J. Brinkhoff and A. E. Parker, "Charge trapping and intermodulation in hemts," in 2004 IEEE MTT-S International Microwave Symposium Digest (IEEE Cat. No. 04CH37535), vol. 2. IEEE, 2004, pp. 799–802.

[Chi35] H. Chireix, "High power outphasing modulation," Proc. IRE, vol. 23, pp. 1370–1392, Nov. 1935

[Cox74] D. C. Cox, "Linear amplification with nonlinear components," IEEE Trans. Commun., vol. COM-22, pp. 1942–1945, Dec. 1974.

[CZE+21] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, “Structured massive access for scalable cell-free massive MIMO systems,” IEEE J. Sel. Areas Commun., vol. 39, no. 4, pp. 1086-1100, Apr. 2021.

[DKH21] 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.

[DHK23] M. F. De Guzman, K. Haneda and P. Kyösti, “Measurement-based MIMO channel model at 140GHz (Version 1),” available: https://doi.org/10.5281/zenodo.7640353.

[DPS20] Erik Dahlman, Stefan Parkvall, Johan Skold, 5G NR - The Next Generation Wireless Access Technology, 2nd Edition, 1. Chapter 26, §26.2: “mmW LO generation and phase noise aspects” - September 18, 2020 ISBN: 9780128223208

[DVB-S2X] ETSI EN 302 307 V1.2.1; Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications. DVB-S2 Extensions (DVB-S2X)

[FDB+19] 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, Pacific Grove, CA, USA, Nov. 2019, pp. 207–214.

[GDS+22] P. R. B Gomes, G T. de Araujo, B. Sokal, A L. F. de Almeida, B. Makki, G. Fodor, “Channel estimation in RIS-assisted MIMO systems operating under imperfections”, [Online]. Available at: https://arxiv.org/pdf/2207.02700.pdf.

[GFF+21] U. Gustavsson, P. Frenger, C. Fager, T. Eriksson, H. Zirath, F. Dielacher, C. Studer, A. Pärssinen, R. Correia, J. Matos, D. Belo, N. Borges Carvalho, “Implementation Challenges and Opportunities in Beyond-5G and 6G Communication,” in IEEE Journal of Microwaves, vol. 1, no. 1, pp. 86-100, winter 2021 (invited).

[GMA+22] H. Guo, B. Makki, M-S. Alouini and T. Svensson, "High-rate uninterrupted internet of vehicle communications in highways: Dynamic blockage avoidance and CSIT acquisition," IEEE Commum. Mag., vol. 60, no. 7, pp. 44-50, Jul. 2022.

[GMM+22] H. Guo, et al. "A Comparison between network-controlled repeaters and reconfigurable intelligent surfaces,” Nov. 2013. [Online]. Available at: https://arxiv.org/abs/2211.06974.

[GS91] A. Ghorbani and M. Sheikhan, “The effect of solid-state power amplifiers (SSPAs) nonlinearities on MPSK and m-QAM signal transmission,” in6th Int. Conf. on Digital Processing of Signals in Communications, Loughborough UK, 1991, p. 193–197.

[GSD+22] F. Gast, M. Schlüter, M. Dörpinghaus, H. Halbauer, and G. Fettweis, “Phase noise tracking for receivers with 1-bit quantization and oversampling,” in ICC 2022 - IEEE International Conference on Communications, May 2022, pp. 1361–1366.

[HEX21-D11] Hexa-X project deliverable D1.1: 6G Vision, use cases and key societal values, Eds. M. Hoffmann, M. Uusitalo, M.-H. Hamon, and B. Richerzhagen, Feb. 2021.

[HEX21-D12] Hexa-X project deliverable D1.2 Expanded 6G vision, use cases and societal values, Eds. M. Hoffmann, M. Uusitalo, M.-H. Hamon, B. Richerzhagen, G. D'Aria, A. Gati and D. Lopez, Apr. 2021.

[HEX21-D12] Hexa-X project deliverable D1.2 Expanded 6G vision, use cases and societal values, Eds. M. Hoffmann, M. Uusitalo, M.-H. Hamon, B. Richerzhagen, G. D'Aria, A. Gati and D. Lopez, Apr. 2021.

[HEX21-D13] Hexa-X project deliverable D1.3: Targets and requirements for 6G - initial E2E architecture, Eds. B.M. Khorsandi, M. Hoffmann, M. Uusitalo, M.-H. Hamon, B. Richerzhagen, G. D'Aria, A. Gati, and D. Lopez, Feb. 2022.

[HEX21-D21] Hexa-X project deliverable D2.1 Towards Tbps Communications in 6G: Use Cases and Gap Analysis, Eds. K. Roth, N. Michailow, A. Pärssinen and H. Halbauer, Jun. 2021.

[HEX21-D22] Hexa-X project deliverable D2.2 Initial radio models and analysis towards ultra-high data rate links in 6G, Eds. Marko E. Leinonen, Dec. 2021.

[HEX21-D31] Hexa-X project deliverable D3.1: Localisation and sensing use cases and gap analysis, Eds. H. Wymeersch, Dec. 2021.

[HEX23-D33] Hexa-X project deliverable D3.3: Final models and measurements for localisation and sensing, Jul. 2023.

[HEX21-D51] Hexa-X project deliverable D5.1: Initial B5G/6G Architectural Components and Enablers, Eds. M. Ericson, H. Flinck, P. Vlacheas, and S. Wänstedt, Dec. 2021.

[HH97] M. Honkanen and S.-G. Haggman, "New aspects on nonlinear power amplifier modeling in radio communication system simulations," in Proceedings of 8th International Symposium on Personal, Indoor and Mobile Radio Communications - PIMRC '97, vol. 3, 1997, pp. 844-848 vol.3.

[HNH+07] A. Hati et al., "Vibration Sensitivity of Microwave Components," 2007 IEEE International Frequency Control Symposium Joint with the 21st European Frequency and Time Forum, Geneva, Switzerland, 2007, pp. 541-546, doi: 10.1109/FREQ.2007.4319131.

[HNH09] A. Hati, C. W. Nelson and D. A. Howe, "Vibration-induced PM and AM noise in microwave components," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 56, no. 10, pp. 2050-2059, October 2009, doi: 10.1109/TUFFC.2009.1288.

[HW21] H. Halbauer, T. Wild, "Towards Power Efficient 6G Sub-THz Transmission," Joint EuCNC & 6G Summit, June 8-11 2021, Porto, Portugal

[IBS+98] M. Ibnkahla, N. Bershad, J. Sombrin, and F. Castanie, "Neural network modeling and identification of nonlinear channels with memory: algorithms, applications, and analytic models," IEEE Transactions on Signal Processing, vol. 46, no. 5, pp. 1208–1220, 1998.

[IEE09] IEEE, "60 GHz Impairment Modeling, Doc. 802.11-09/1213r1," IEEE Std 802.11-09, pp. 1–29, Nov 2009.

[IEE16] IEEE, "IEEE TGay Evaluation Methodology, Doc. 802.11-15/0866r2," IEEE Std 802.11-15, pp. 1–27, Oct 2016.

[Imm04] K. A. Schouhamer Immink, Codes for Mass Data Storage Systems, 2. ed., Shannon Foundation Publishers, 2004.

[ITU P.2040-2] ITU-R recommendation P.2040-2, “Effects of building materials and structures on radiowave propagation above about 100 MHz”, Jul. 2021.

[KGH+22] 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, Jan. 2022.

[KKP+14] M. R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, “Calculation of the performance of communication systems from measured oscillator phase noise,” IEEE Trans. Circuits Syst. I, vol. 61, no. 5, pp. 1553–1565, May 2014.

[LDF15] L. Landau, M. Dörpinghaus, and G. Fettweis, “Communications employing 1-bit quantization and oversampling at the receiver: faster-than-Nyquist signaling and sequence design,” International Conference on Ultra Wideband (ICUWB), Montreal, Canada, Oct. 2015

[LJT+19] 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 Transactions on Microwave Theory and Techniques, vol. 67, no. 12, pp. 5282-5295, Dec. 2019, doi: 10.1109/TMTT.2019.2951773.

[LMN+22] M. E. Leinonen, J.-P. Mäkelä, K. Nevala, N. Tervo, and A. Pärssinen, “Repeatability of 220 - 330 GHz variable waveguide attenuator with frequency extenders for 6G measurements,” in 2022 98th ARFTG Microwave Measurement Conference (ARFTG), 2022, pp. 1–4

[LNT+21] M. E. Leinonen, K. Nevala, N. Tervo, and A. Pärssinen, “Linearity measurement of 6g receiver with one transmission frequency extender operating at 330 GHz,” in 2021 96th ARFTG Microwave Measurement Conference (ARFTG), 2021, pp. 1–4.

[LR11] C. Liang and B. Razavi, "Transmitter Linearization by Beamforming," in IEEE Journal of Solid-State Circuits, vol. 46, no. 9, pp. 1956-1969, Sept. 2011, doi: 10.1109/JSSC.2011.2148530.

[LTJ+19] M. E. Leinonen, N. Tervo, M. Jokinen, O. Kursu and A. Pärssinen, "5G mm-Wave Link Range Estimation Based on Over-the-Air Measured System EVM Performance," 2019 IEEE MTT-S International Microwave Symposium (IMS), Boston, MA, USA, 2019, pp. 476-479

[LWZ20] Y. Li, X. Wang and A. Zhu, "Sampling Rate Reduction for Digital Predistortion of Broadband RF Power Amplifiers," in IEEE Transactions on Microwave Theory and Techniques, vol. 68, no. 3, pp. 1054-1064, March 2020, doi: 10.1109/TMTT.2019.2944813.

[LXA+21] V. Lampu et al., "Energy-Efficient Array Transmitters Through Outphasing and Over-the-Air Combining," 2021 55th Asilomar Conference on Signals, Systems, and Computers, 2021, pp. 608-615.

[MMC+20] C. Madapatha, B. Makki, C. Fang, et al., “On integrated access and backhaul networks: Current status and potentials,” IEEE open j. Commun. Soc., vol. 1, pp. 1374–1389, Sep. 2020.

[MMH+22] C. Madapatha, B. Makki, H. Guo, and T. Svensson., Constrained deployment optimization in integrated access and backhaul networks. arXiv preprint arXiv:2210.05253.

[MMM+21] 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,” IEEE Open j. Commun. Soc., vol. 2, pp. 2273–2291, 2021

[MSH+18] S. Mondal, R. Singh, A. I. Hussein and J. Paramesh, "A 25–30 GHz Fully-Connected Hybrid Beamforming Receiver for MIMO Communication," in IEEE Journal of Solid-State Circuits, vol. 53, no. 5, pp. 1275-1287, May 2018.

[MST+02] 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.

[Mur22] B. Murmann, "ADC Performance Survey 1997-2022," [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html , (version dated 22020719).

[NAY+17] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-free massive MIMO versus small cells,” IEEE Trans. Wireless Commun., vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[NJK+18] 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," 12th European Conference on Antennas and Propagation (EuCAP 2018), London, UK, 2018.

[NSJ+21] P. Neuhaus, M. Schlüter, C. Jans, M. Dörpinghaus, and G. Fettweis, “Enabling energy-efficient Tbit/s communications by 1-bit quantization and oversampling,” in Proc. 2021 Joint European Conf. on Networks and Communications 6G Summit (EuCNC/6G Summit), Jun. 2021, pp. 84–89.

[OML09] M. O’Droma, S. Meza, and Y. Lei, “New modified Saleh models for memoryless nonlinear power amplifier behavioural modelling,” IEEE Communications Letters, vol. 13, no. 6, pp. 399–401, 2009.

[PAB+20] A. Pärssinen, M. Alouini, M. Berg, T. Kuerner, P. Kyösti, M.E. Leinonen, M. Matinmikko-Blue, E. McCune, U. Pfeiffer, P. Wambacq, (Eds.). (2020). White Paper on RF Enabling 6G – Opportunities and Challenges from Technology to Spectrum [White paper]. (6G Research Visions, No. 13). University of Oulu. http://urn.fi/urn:isbn:9789526228419

[PM05] J. Pedro and S. Maas, “A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches,” IEEE Transactions on Microwave Theory and Techniques, vol. 53, no. 4, pp. 1150-1163, 2005.

[Puerta2022] R. Puerta, O. Ozolins, A. Djupsjöbacka, V. Bobrovs, S. Popov, and X. Pang, ‘5G/NR Conformance Testing of Analog Radio-over-fiber Fronthaul Links’, in *2022 IEEE International Topical Meeting on Microwave Photonics (MWP)*, 2022, pp. 1–4.

[Rap91] C. Rapp, “Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal,” in ESA Special Publication, ser. ESA Special Publication, B. Kaldeich, Ed., vol. 332, Oct. 1991, pp. 179–184.

[REA+21] N. Rajatheva et al., “Scoring the terabit/s goal: Broadband connectivity in 6G,” Feb. 2021. [Online]. Available: https://arxiv.org/pdf/2008.07220.pdf.

[RFWL22] “Rf weblab,” http://dpdcompetition.com/rfweblab/measurement-setup/, accessed: 2022-11-11.

[RKL+20] K. Rikkinen, P. Kyösti, M. E. Leinonen, M. Berg and A. Pärssinen, "THz Radio Communication: Link Budget Analysis toward 6G," in IEEE Communications Magazine, vol. 58, no. 11, pp. 22-27, November 2020.

[Rot06] R. Roth, Introduction to Coding Theory. Cambridge University Press, 2006.

[SDF22] M. Schlüter, M. Dörpinghaus, and G. Fettweis, "Joint phase and timing estimation with 1-bit quantization and oversampling," IEEE Trans. Commun., vol. 70, no. 1, pp. 71–86, Jan. 2022.

[SHS+19] S. Shahramian, M. J. Holyoak, A. Singh and Y. Baeyens, "A Fully Integrated 384-Element, 16-Tile, W -Band Phased Array With Self-Alignment and Self-Test," in IEEE Journal of Solid-State Circuits, vol. 54, no. 9, pp. 2419-2434, Sept. 2019.

[SNL+21] D. Selimis, K. Ntontin, F. Lazarakis and K. Haneda, "Initial Investigation of D-band Small-Scale Fading Statistics," *2021 15th European Conference on Antennas and Propagation (EuCAP)*, 2021, pp. 1-5.

[SOG+08] D. Schreurs, M. O'Droma, A. A. Goacher, and M. Gadringer, RF power amplifier behavioral modeling. Cambridge university press New York, NY, USA:, 2008, vol. 1.

[SRL+21] S. P. Singh, T. Rahkonen, M. E. Leinonen, and A. Pärssinen, "A 290 GHz low noise amplifier operating above fmax/2 in 130 nm SiGe technology for sub-THz/THz receivers," in 2021 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), 2021, pp. 223–226.

[ST05] R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters. New York: Wiley, 2005.

[TAT+16] N. Tervo, J. Aikio, T. Tuovinen, T. Rahkonen and A. Parssinen, "Effects of PA Nonlinearity and Dynamic Range in Spatially Multiplexed Precoded MIMO Systems," *European Wireless 2016; 22th European Wireless Conference*, Oulu, Finland, 2016, pp. 1-6.

[TAT+17] N. Tervo, J. Aikio, T. Tuovinen, T. Rahkonen and A. Parssinen, "Digital predistortion of amplitude varying phased array utilising over-the-air combining," 2017 IEEE MTT-S International Microwave Symposium (IMS), Honololu, HI, USA, 2017, pp. 1165-1168, doi: 10.1109/MWSYM.2017.8058809.

[Ter22] N. Tervo, "Concepts for radiated nonlinear distortion and spatial linearization in millimeter-wave phased arrays", PhD Thesis, Oulu, 2022

[TLM19] Texas Instruments, "LMX2594 15-GHz Wideband PLLATINUM™ RF Synthesizer With Phase Synchronization and JESD204B Support". Product datasheet. Assessed Feb 2023. (online) https://www.ti.com/lit/ds/symlink/lmx2594.pdf?

[TR 37.910] 3GPP TR 37.910 v16.1.0, "Study on Self Evaluation Towards IMT-2020 Submission (Release 16)"

[TTP17] T. Tuovinen, N. Tervo and A. Pärssinen, "Analyzing 5G RF System Performance and Relation to Link Budget for Directive MIMO," in *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6636-6645, Dec. 2017, doi: 10.1109/TAP.2017.2756848.

[VRD03] J. Vuolevi and T. Rahkonen, Distortion in RF power amplifiers. Artech house, 2003.

[WAL99] R. H. Walden, “Analog-to-digital converter survey and analysis”, IEEE J. Select. Areas Commun., vol. 17, pp. 539-550, April 1999.

[WBJ03] G. White, A. Burr, and T. Javornik, “Modelling of nonlinear distortion in broadband fixed wireless access systems,” Electronics Letters, vol. 39, pp. 686–687(1), April 2003

[WWH+23] Power Amplifiers Performance Survey 2000-Present,” [Online]. Available: https://gems.ece.gatech.edu/PA survey.html

[XWW+21] Z. Xing, R. Wang, J. Wu, and E. Liu, “Achievable rate analysis and phase shift optimization on intelligent reflecting surface with hardware impairments,” IEEE Trans. Wireless Commun., vol. 20, no. 9, pp. 5514– 5530, Sept. 2021.

[YMY+19] X. Yang, M. Matthaiou, J. Yang, C.-K. Wen, F. Gao, and S. Jin, “Hardware-constrained millimeter-wave systems for 5G: challenges, opportunities, and solutions,” IEEE Communications Magazine, vol. 57, no. 1, pp. 44–50, Jan. 2019.

[YYC22] H. Yu, X. Yi, and G. Caire, “Topological pilot assignment in large-scale distributed MIMO networks,” IEEE Trans. Wireless Commun., vol. 21, no. 8, pp. 6141–6155, 2022.

[ZDS+21] A. Zappone, M. Di Renzo, F. Shams, X. Qian, and M. Debbah, “Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments,” IEEE Trans. Wireless Commun., vol. 20, no. 1, pp. 126–141, Jan. 2021.

[ZGD+22] S. Zeitz, F. Gast, M. Dörpinghaus, and G. Fettweis, “On the Bayesian Cramér-Rao bound for phase noise estimation based on 1-bit quantized samples,” in IEEE Global Communications Conference (GLOBECOM), Rio de Janeiro, Brazil, 2022.

[ZKB+23] P. Zhang, P. Kyösti, M. Bengtson, V. Hovinen, K. Nevala, J. Kokkoniemi, A. Pärssinen, ”Measurement-Based Characterization of D-Band Human Body Shadowing,” 17th European Conference on Antennas and Propagation (EuCAP 2023), March 2023 (accepted).