



Politechnika Wroclawska

FIELD OF SCIENCE: Engineering and Technology

DISCIPLINE OF SCIENCE: Information and Communication Technology

DOCTORAL DISSERTATION

Improvements of the Performance of Energy-Efficient 5G Massive MIMO Base Stations

Łukasz Skomra

Supervisor:

Prof. dr hab. inż. Kamil Staniec

Assistant supervisor:

Dr.-Ing. Björn Jelonnek

Keywords: Massive MIMO, distortion, energy consumption, peak-to-average-power ratio reduction

WROCŁAW 2024

Abstract

Massive MIMO (multiple-input multiple-output) base stations, deployed in 5G cellular networks, enable unprecedented increase in the achievable data rates due to spatial multiplexing of multiple users. Since the base station sites account for most of the electric energy consumed in the cellular networks, an energy-efficient operation of Massive MIMO base stations is the key for sustainable provision of high-end performance to many users.

A considerable portion of the Massive MIMO literature either assumes that the distortion introduced in the base station transmitters (TXs) combines incoherently at the user equipment (UE) or ignores the impact of the hardware non-idealities altogether. This may be a feasible simplification when saturation in the power amplifiers (PAs) is avoided. However, usage of a high input back-off (IBO) to achieve this degrades the energy efficiency. Aiming to decrease the IBO, many studies on peak-to-average-power ratio (PAPR) reduction exploit the excess degrees of freedom in the Massive MIMO channel, striving to achieve a distortion-free transmission. This constraint may be too restrictive in practice. Receiver noise and intra- or inter-cell interference experienced by the UEs limits downlink (DL) data rates, regardless of any distortion effects.

The ultimate objective of this dissertation is to clarify whether the residual TX impairments in Massive MIMO base stations have substantial impact on the DL data rates. The nonlinear distortion introduced by peak power limitation, resulting either from saturation in linearized PAs or PAPR reduction, is in the focus. The analysis covers compact arrays, utilizing either linear or rectangular layout of co-located antennas.

The first contribution of this dissertation is the proposal of two models for predicting the per user error vector magnitude (EVM) performance in line-of-sight (LoS) scenarios, based on the expectation of the radiation patterns of third-order intermodulation products (IMPs). The proposed simplified best-case model relies on categorization of IMPs, as either user-directed or spatially filtered (radiating in distinct directions). The portion of the IMP energy explicitly directed to users reduces with increasing the number of spatially multiplexed UEs. However, analytical analysis revealed existence of unfavorable combinations of UE angular locations, resulting in part of the spatially filtered IMPs to radiate in the UE directions as well. The proposed precoding-based model covers the cases of unfavorable combinations of angular directions and reflects the effects of limited angular resolution of the antenna array.

The comparison of the EVM statistics obtained from Monte Carlo trials of varying UE locations for the numerical simulation with precoding-based model and the link-level

simulation (LLS), using iterative clipping and filtering (ICF) as the source of nonlinear distortion, suggests that the proposed model can relatively well reflect the performance of peak power limitation. The prediction was generally pessimistic, as expected due to modeling only third-order IMPs, with a typical deviation of about 0.7 dB.

The spectral efficiency calculation was extended by reflecting the power of distortion received by UEs (estimated based on the proposed EVM prediction models), to consider the impact of the residual TX impairments on the achievable DL data rates. The relation between the signal-to-interference-plus-noise ratio (SINR) distribution in a multi-cell network and the impact of distortion was studied using Monte Carlo numerical simulations. As expected, the range of observed SINRs was relatively wide (about 20 dB spread between 10th and 90th percentiles), leading to different impact for different UEs. In the 16-cell example with ICF, decrease of the IBO from 10 dB to 6 dB resulted in over 20% reduction in the estimated energy consumption at a cost of less than 3% drop in the cell throughput.

Worst-case combinations of UE angular locations, resulting in all the IMPs to radiate in the UE directions, were identified analytically and confirmed in the LLS. In these outlier scenarios a full coherent combination of distortion at the UE locations is expected, irrespective of the number of antennas and spatially multiplexed UEs, resulting in lack of observable improvement in the per user EVM compared to the per antenna EVM. For example, for 6 spatially multiplexed UEs about 6 dB difference in the per user EVM between the best- and worst-case sets of user angular directions was observed in LLS. It was demonstrated that precoding-based prediction of the radiation pattern of spatially filtered IMPs can be used to improve the EVM performance by avoiding co-scheduling certain UEs.

The second contribution of this dissertation is a novel approach for exploiting the excess degrees of freedom available in the Massive MIMO channel, by using the power headroom in a subset of antennas to compensate for the distortion introduced in a different subset of antennas. The efficacy of the proposed solution was demonstrated in LLS for the ICF-based PAPR reduction as the source of distortion in a tapered linear array. An improvement of over 10 dB in EVM was observed in all the verified cases with LoS, when half of the antennas were compensating and zero forcing was used to precode the compensation signal.

This dissertation proves that the TX distortion, introduced in Massive MIMO base stations, may have visible impact on the DL data rates. Since a distortion-free operation is not required in practice, the balance between the DL data rates and energy consumption can be explored, taking into consideration the SINR distribution expected in a given deployment.

Streszczenie

Stacje bazowe wykorzystujące systemy wieloantenowe Massive MIMO (ang. *Massive multiple-input multiple-output*), używane w sieciach komórkowych 5. generacji, umożliwiają bezprecedensowy wzrost osiągalnych szybkości transmisji danych dzięki multipleksowaniu przestrzennemu wielu użytkowników. Ponieważ obiekty infrastruktury stacji bazowych odpowiadają za większość zużycia energii elektrycznej w sieciach komórkowych, energooszczędna praca stacji bazowych Massive MIMO jest kluczem do zrównoważonego zapewnienia wysokiej wydajności wielu użytkownikom sieci.

Znaczna część literatury dotyczącej Massive MIMO albo zakłada, że zniekształcenia wprowadzane w nadajnikach (ang. *transmitters*, TXs) stacji bazowych łączą się w sposób niespójny po stronie użytkownika, albo wprost ignoruje w analizie wpływ niedoskonałości sprzętowych. Tego rodzaju uproszczenia mogą prowadzić do użytecznych w praktyce wniosków, pod warunkiem unikania nasycenia we wzmacniaczach mocy (ang. *power amplifiers*, PAs). Jednak zastosowanie w tym celu restrykcyjnego ograniczenia poziomu mocy wejściowej (ang. *input back-off*, IBO) pogarsza efektywność energetyczną. Dążąc do zmniejszenia wymaganego IBO, wiele prac badawczych nad redukcją stosunku mocy szczytowej do średniej (ang. *peak-to-average-power ratio*, PAPR) wykorzystuje stopnie swobody w kanale Massive MIMO dążąc do osiągnięcia transmisji wolnej od zniekształceń. W praktyce może to być nadmiernie restrykcyjne założenie. Zarówno szum odbiornika jak i odebrane zakłócenia z transmisji do innych użytkowników (w tej samej lub innej komórce) ograniczają szybkość łącza w dół (ang. *downlink*, DL), niezależnie od wpływu zniekształceń.

Celem niniejszej rozprawy jest wyjaśnienie, czy zniekształcenia wprowadzane do sygnału w stacjach bazowych Massive MIMO mogą istotnie wpływać na szybkość transmisji danych w DL. Zniekształcenia nieliniowe wprowadzane przez ograniczenie mocy szczytowej, wynikające z nasycenia w PA lub z redukcji PAPR, stanowią kluczowy element analizy. Zakres badań obejmuje kompaktowe liniowe i prostokątne układy antenowe.

Pierwszym istotnym rezultatem rozprawy jest zaproponowanie dwóch modeli predykcyjnych dla określenia wielkości wektora błędu (ang. *error vector magnitude*, EVM) po stronie użytkownika, w scenariuszach komunikacji z LoS (ang. *line-of-sight*), w oparciu o oczekiwania dotyczące charakterystyki promieniowania produktów intermodulacji (ang. *intermodulation products*, IMPs) trzeciego rzędu. Pierwszy z zaproponowanych modeli bazuje na kategoryzacji IMPs na kierowane do użytkowników i filtrowane przestrzennie (promieniujące w konkretnych kierunkach). Część energii IMP, która jest wprost kierowana

w kierunkach zgodnych z sygnałem transmitowanym do użytkowników zmniejsza się wraz ze wzrostem liczby multipleksowanych przestrzennie użytkowników. Jednakże wykazano analitycznie, że istnieją niekorzystne kombinacje kierunków kątowych użytkowników, skutkujące promieniowaniem części filtrowanych przestrzennie IMPs również w kierunkach użytkowników. Drugi z modeli, bazujący na analizie macierzy prekodowania, obejmuje przypadki niekorzystnych kombinacji kierunków kątowych transmisji i odzwierciedla skutki ograniczonej rozdzielczości kątowej układu antenowego.

Badania zależności EVM od kombinacji lokalizacji użytkowników przeprowadzono w oparciu o metodę Monte Carlo. Porównano statystyki uzyskane dla symulacji numerycznych z modelu predykcyjnego opartego na prekodowaniu i symulacji poziomu łącza (ang. *link-level simulation*, LLS), wykorzystując redukcję PAPR z podejściem opartym o ICF (ang. *iterative clipping and filtering*) jako źródło zniekształceń nieliniowych. Uzyskana prognoza była ogólnie pesymistyczna, zgodnie z oczekiwaniami wynikającymi z modelowania jedynie IMP trzeciego rzędu, z typowym odchyleniem około 0.7 dB. Wynik ten sugeruje, że proponowany model może stosunkowo dobrze odzwierciedlać zależność od kierunków kątowych transmisji w statystykach EVM, dla zniekształceń nieliniowych wynikających z ograniczenia mocy szczytowej.

W rozważaniach wpływu zniekształceń na osiągalne szybkości transmisji danych w DL, obliczenia wydajności widmowej uzupełniono o czynnik uwzględniający moc zniekształceń odbieranych przez użytkownika, oszacowaną na podstawie modeli predykcyjnych EVM. Związek między rozkładem stosunku sygnału do zakłóceń i szumu (ang. *signal-to-interference-plus-noise ratio*, SINR) w sieci komórkowej a wpływem zniekształceń badano za pomocą symulacji numerycznych wykorzystując metodę Monte Carlo. Zgodnie z oczekiwaniami, zakres zaobserwowanego SINR był stosunkowo szeroki (około 20 dB rozpiętości pomiędzy 10. a 90. percentylem), co skutkowało różnym poziomem istotności wpływu zniekształceń na poszczególne transmisje. W przykładzie z 16 komórkami, obniżenie IBO z 10 dB do 6 dB spowodowało ponad 20% redukcję szacowanego zużycia energii, kosztem spadku przepustowości komórek o mniej niż 3%.

Specyficzne zestawy kombinacji kierunków kątowych transmisji prowadzące do promieniowania wszystkich wygenerowanych IMPs w kierunku użytkowników zostały zidentyfikowane analitycznie i potwierdzone w LLS. W tych odstających scenariuszach można oczekiwać pełnej spójnej kombinacji zniekształceń w lokalizacjach użytkowników, niezależnie od liczby anten oraz ilości zmultipleksowanych przestrzennie warstw transmisji. W efekcie, widoczny jest brak poprawy EVM mierzonego dla użytkownika w porównaniu

z EVM mierzonym dla anteny nadawczej. Przykładowo, dla 6 zmultipleksowanych przestrzennie użytkowników zaobserwowano w LLS około 6 dB różnicy w EVM pomiędzy najlepszymi i najgorszymi zestawami kierunków kątowych. Wykazano, że oczekiwana charakterystyka promieniowania IMPs (określona na podstawie modelu opartego o macierz prekodowania) może być wykorzystana do poprawy EVM. Efekt ten został uzyskany poprzez unikanie przydziału zasobów w danym przedziale czasu użytkownikom, którzy spowodowaliby znaczący wzrost części mocy filtrowanych przestrzennie IMPs promieniowanej do użytkowników.

Drugim z istotnych rezultatów niniejszej rozprawy jest propozycja nowatorskiego podejście do wykorzystania nadmiarowych stopni swobody dostępnych w kanale Massive MIMO, poprzez wykorzystanie zapasu mocy w podzbiorze anten do kompensacji zniekształceń wprowadzanych w innym podzbiorze anten. Skuteczność proponowanego rozwiązania wykazano w LLS, w oparciu o wykorzystanie redukcji PAPR opartej na ICF, jako źródła zniekształceń w taperowanym liniowym układzie antenowym. Zaobserwowano poprawę EVM o ponad 10 dB we wszystkich zweryfikowanych przypadkach z LoS, gdy do kompensacji wykorzystano połowę anten, a sygnał kompensacyjny był prekodowany przy pomocy podejścia opartego na wymuszaniu zer.

Niniejsza rozprawa dowodzi, że zniekształcenia TX wprowadzane w stacjach bazowych Massive MIMO mogą mieć istotny wpływ na ograniczenie szybkości transmisji danych w DL. Ponieważ w praktyce nie jest wymagana praca bez zniekształceń, poprawa równowagi między szybkością transmisji danych w DL a zużyciem energii może zostać zrealizowana w oparciu o rozkład SINR w komórce, oczekiwany w danym przypadku użycia stacji bazowej.

Acknowledgements

I would like to express my deepest gratitude and appreciation to my supervisors. For their guidance, as well as their time and energy spent in supporting me during the doctoral research journey. Professor Kamil Staniec, specifically for coaching, mentoring, and essential support in navigating the intricacies of the scientific research universe. Doctor Björn Jellonek, Nokia Bell Labs Fellow, particularly for encouraging me to step into this path, his open and supporting attitude along the way, and all the discussion and feedback on plausibility of the ideas and results.

This endeavor would not have been possible without the support from the management team at Nokia. Thank you for enabling me to grow as a researcher.

The “Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency” book has proven to be an invaluable source of well-structured information, providing a solid foundation for my doctoral research in the area of Massive MIMO. I would like to express my deepest appreciation to its authors: Emil Björnson, Jakob Hoydis, and Luca Sanguinetti.

I would like to extend my gratitude to the Polish Ministry of Science and Higher Education for establishing the “Doktorat Wdrożeniowy” program and deciding to fund this research project under its framework. Many thanks to the staff of the Doctoral School for facilitating the process. I would also like to recognize the support provided by the Wrocław Centre for Networking and Supercomputing by allocating computational resources for my research project.

Special thanks to my colleagues from Nokia, for all the valuable interactions. I am deeply indebted to Gunter and Hartmut, who strongly influenced my attitude for doing research and communicating its results, especially in the beginning of my career at Nokia.

Many thanks to my family, for believing in me. I am especially grateful to my parents, for instilling in me a sense of curiosity and a passion for knowledge.

Finally, I would like to thank the most important person in my life, my wife Ania, who has stood by me every step of the way. Offering her unconditional support and fueling my perseverance. Thank you.

I appreciate all the positive energy, words of encouragement, and offers of support I have received from many who are not explicitly mentioned here by name. My four-year PhD research journey culminates with the doctoral dissertation which I hereby proudly present to you.

Contents

Abstract	i
Acknowledgements	vi
Abbreviations, Symbols, and Notation	xi
1. Introduction.....	1
1.1. Research Aim and Objectives.....	2
1.2. Theses	2
1.3. Overview of the Structure and Contributions of the Dissertation	3
1.4. List of Prepared Publications and Patent Application	4
2. Achievable Data Rates and Energy Consumption as Key Performance Indicators	6
2.1. Achievable Downlink User Data Rates and Cell Throughput.....	6
2.2. Base Station Energy Consumption in Downlink Transmission	8
3. State-of-the-Art	10
3.1. General Background	10
3.2. Residual Transmitter Impairments in Massive MIMO Base Stations.....	12
3.3. Peak-to-Average-Power Ratio Reduction in Massive MIMO.....	18
3.4. Cognitive Gaps and Associated Theses.....	25
4. Predicting the In-Band Impact of the Residual Transmitter Impairments	27
4.1. System Model.....	28
4.1.1. Allocation of Power and Frequency Resources.....	29
4.2. Models of the Residual Transmitter Impairments	32
4.2.1. Nonlinear Distortion Introduced by Peak Limitation	32
4.2.2. Noise Added by Quantization	33
4.3. Preliminaries.....	34
4.3.1. Tolerable Error Vector Magnitude	34

4.3.2.	Beam Steering and Array Factor	35
4.3.3.	Radiated Intermodulation Products	37
4.4.	Proposed Error Vector Magnitude Performance Prediction Models for Nonlinear Distortion in Line-of-Sight Transmission.....	39
4.4.1.	Per Antenna Error.....	40
4.4.2.	Per User Error.....	42
4.5.	Expected Error Vector Magnitude Performance for Quantization.....	54
4.6.	Methods	55
4.6.1.	Antennas	55
4.6.2.	Resource Allocation	56
4.6.3.	Channels and Precoding	56
4.6.4.	Equalization and Calculation of the Error Vector Magnitude.....	57
4.6.5.	Repeatability of Simulation Results	57
4.6.6.	Rejection of Selected Simulation Results.....	57
4.7.	Simulation Results and Discussion	57
4.7.1.	Accuracy of the per Antenna Error Model for Peak Power Limitation ..	58
4.7.2.	Best- and Worst-Cases for Line-of-Sight Channels	60
4.7.3.	Statistics of the Error Level for Randomly Placed Users.....	64
4.7.4.	Performance in Link-Level Simulation for Non-Line-of-Sight Channels 73	
4.7.5.	Impact of the Power Allocation Strategy	74
4.8.	Summary and Proof of Thesis	78
5.	Impact of Distortion on Downlink Data Rates and Energy Consumption	81
5.1.	Preliminaries.....	81
5.1.1.	Practical Limitations of the Achievable Data Rates.....	82
5.1.2.	Wideband Signal-to-Noise Ratio.....	84
5.1.3.	Channel Hardening.....	86
5.2.	Methods	87

5.2.1.	Reflecting the Impact of Nonlinear Distortion in a Multi-Cell Network	88
5.3.	Simulation Results and Discussion	88
5.3.1.	Block Error Rate Performance with Transmitter and Receiver Noise	89
5.3.2.	Data Rate Impact of Unequal Power Allocation	91
5.3.3.	Balance Between the Achievable Data Rates and Energy Consumption in Multi-Cell Networks	93
5.3.4.	Performance Improvement Potential of Distortion-Aware Scheduling	98
5.4.	Summary and Proof of Thesis	107
6.	Exploiting Spatial Degrees of Freedom to Compensate Transmitter Distortion	110
6.1.	System Model	111
6.2.	Preliminaries	111
6.2.1.	Using Reserved Antennas to Compensate Transmitter Distortion	111
6.2.2.	Per Antenna Power in a Tapered Array	112
6.3.	Proposed Compensation Method for Signal Distortion Introduced by Peak-to-Average-Power Ratio Reduction	114
6.3.1.	Compensating Distortion from a Subset of Antennas	115
6.3.2.	Calculation of the Compensation Signals	118
6.4.	Methods	118
6.5.	Simulation Results and Discussion	119
6.5.1.	Reduction of the Required Peak Power Capability	119
6.5.2.	Improvement in the Level of Error at the Users	121
6.5.3.	Impact on the Energy Consumption	124
6.6.	Summary and Proof of Thesis	127
7.	Conclusions	128
	Bibliography	131
	List of Figures	140
	List of Tables	146
A.	Antenna Arrays	147

A.1.	Uniform Linear Array with Horizontal Steering	148
A.1.1.	Geometry, Orientation and Beam Steering	148
A.1.2.	Array Factor.....	150
A.1.3.	Critically Spaced Array	151
A.2.	Uniform Rectangular Array.....	153
A.2.1.	Geometry, Orientation and Beam Steering	153
A.2.2.	Array Factor.....	156
A.2.3.	Critically Spaced Array	157
A.3.	Beamforming.....	157
B.	Intermodulation Beams.....	159
B.1.	Intermodulation Products	159
B.2.	Intermodulation Distortion Radiated from Antenna Array	164
B.2.1.	Example for Uniform Linear Array.....	165
C.	The Link-Level Simulation Environment.....	169
C.1.	Structure and Operation.....	169
C.2.	Calibration	170
D.	Simulation Results Supplementing the Core Analysis	172
D.1.	Higher-Order Intermodulation Products.....	172
D.2.	Angular Positions of Users in Non-Line-of-Sight Channels and Power Variation Over Antennas	173
D.3.	Statistics of the Error Level for Hard Clipping	176
D.4.	Quantization Noise Radiated Over Multi-Path Channels.....	177

Abbreviations, Symbols, and Notation

Abbreviations

3GPP	3rd Generation Partnership Project
5G	5th Generation
AF	Array Factor
AWGN	Additive White Gaussian Noise
BLER	Block Error Rate
CCDF	Complementary CDF
CDF	Cumulative Distribution Function
CDL	Clustered Delay Line
CL	Clipping Level
CP-OFDM	Cyclic Prefix OFDM
CR	Clipping Ratio
CSI	Channel State Information
DAC	Digital-to-Analog Converter
DL	Downlink
DMRS	Demodulation Reference Signal
DS	Delay Spread
EVM	Error Vector Magnitude
FDD	Frequency-Division Duplex
FDM	Frequency-Division Multiplexing
HC	Hard Clipping
HW	Hardware
i.i.d.	independent and identically distributed
IBO	Input Back-Off
ICF	Iterative Clipping and Filtering
IM	Intermodulation
IMD	Intermodulation Distortion
IMP	Intermodulation Product
IUI	Inter-User Interference
JPP	Joint Precoding and PAPR reduction
LLS	Link-Level Simulation (or Simulator)
LoS	Line-of-Sight
MCS	Modulation and Coding Scheme
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output

MMSE	Minimum Mean-Squared Error
MU-MIMO	Multi-User MIMO
M-MMSE	Multicell MMSE
NLoS	Non-Line-of-Sight
NR	New Radio
OFDM	Orthogonal Frequency-Division Multiplexing
OOB	Out-Of-Band
OTA	Over-The-Air
PA	Power Amplifier
PAPC	Per Antenna Power Constraint
PAPR	Peak-to-Average-Power Ratio
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RE	Resource Element
RMS	Root Mean Square
RRM	Radio Resource Management
RX	Receiver or Receive
RZF	Regularized Zero Forcing
SC	Subcarrier
SDR	Signal-to-Distortion Ratio
SE	Spectral Efficiency
SER	Symbol Error Rate
SINDR	Signal-to-Interference-plus-Noise-plus-Distortion Ratio
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
SNDR	Signal-to-Noise-plus-Distortion Ratio
SPC	Sum Power Constraint
SU-MIMO	Single-User MIMO
TDD	Time-Division Duplex
TX	Transmitter or Transmit
UE	User Equipment
UL	Uplink
ULA	Uniform Linear Array
UMa	Urban Macro
UMi	Urban Micro
URA	Uniform Rectangular Array
ZF	Zero Forcing

Symbols

d	distortion signal
D	power of received distortion
I	power of received interference
J	number of SCs, indexed with j
K	number of UEs, indexed with k
M	number of antennas in the array, indexed with m
M_y	number of antennas in a single row (along the y -axis) of the array
M_z	number of antennas in a single column (along the z -axis) of the array
N	power of receiver noise
P	sum output power
S	power of received wanted data signal
w	precoding weight
x	precoded signal, to be transmitted from the array
y	signal received by the UE
z	signal received by the UE, after equalization
β	average channel gain
$\Delta\gamma$	linear phase gradient over antennas, used for beam steering
γ	phase of a precoding weight w per antenna, used for beamforming
θ	elevation angle
κ	angular wavenumber
λ	linear wavelength
ϕ	azimuth angle
ω	angular frequency

Mathematical notation

Lower-case boldface letters represent column vectors (e.g., \mathbf{x} , \mathbf{w}), upper-case boldface letters denote matrices (e.g., \mathbf{X} , \mathbf{W}), while scalars are represented by italic letters (e.g., x , w , M , P).

The superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, denote complex conjugate, transpose, and conjugate transpose, respectively. Euclidian norm is expressed by $\|\cdot\|$, expectation operator by $\mathbb{E}\{\cdot\}$, and variance operator by $\text{Var}\{\cdot\}$.

The set of complex-valued matrices with K rows and M columns is denoted by $\mathbb{C}^{K \times M}$.

Chapter 1

Introduction

Cellular networks enable seamless wireless connectivity for mobile users. Transition from the aspiration of connecting people to the endeavor of providing both human and machine users with reliable means of high-performance mobile communication, pushes academia and industry to innovate and rethink the mobile networks of the future. The concept of Massive MIMO (multiple-input multiple-output) networks is arguably a pivotal step in that journey.

The capability of Massive MIMO base stations to spatially multiplex multiple users leads to unprecedented enhancement of the spectral efficiency (SE) [1, Sec. 1.3], enabling increase in the achievable peak and average data rates. This allows for rational utilization of the limited spectrum resources. At the same time, substantial portion of the electric energy used by the cellular networks is attributed to the relatively high energy consumption in base stations [1, Sec. 5.1]. By focusing the radiated energy to selected directions, Massive MIMO base stations allow to use lower transmit power compared with conventional base stations. Further reduction of the energy consumption should be possible by improving the energy efficiency of the components of the base station.

It is suggested that the requirements on the quality of base station hardware (HW) can be significantly relaxed when the number of antennas is increased [1, Sec. 6.4]. This could result in an improved energy efficiency, by allowing operation of power amplifiers (PAs) closer to their saturation points [1, Sec. 5.1], while maintaining high data rates. However, the extent of possible relaxation depends on the radiation pattern of distortion, which can be fully aligned with the user direction in the worst case [2].

The ultimate objective of this dissertation is to clarify whether the distortion introduced in the transmitters of practical Massive MIMO base stations, deployable in 5G cellular networks, may noticeably impact the downlink (DL) data rate performance.

While an expanded view on the relation between the achievable DL data rates and the energy consumption of the base station is provided in Chapter 2, the following sections introduce the research aim, objectives, and theses of this dissertation. An overview of the structure and contributions, covered in Section 1.3, is followed by a description of the publications and patent application prepared as part of the research effort covered by this dissertation.

1.1. Research Aim and Objectives

The research aim of this dissertation is to consider how the balance between the achievable DL data rates and the energy consumption of Massive MIMO base stations can be improved.

The scope is on Massive MIMO base stations with co-located (i.e., not distributed) antennas, in urban macro (UMa) and urban micro (UMi) deployments of cellular network. The focus is on transmission using a single 5G new radio (NR) orthogonal frequency-division multiplexing (OFDM) carrier to users with single-antenna user equipment (UE). This analysis covers the impact of the residual transmitter (TX) impairments, in Massive MIMO base station, resulting from: limiter-based peak-to-average-power ratio (PAPR) reduction, linearized PAs, and quantization.

Following research objectives are addressed in this dissertation:

- Verify, in link-level simulation, the impact of the residual TX impairments on the quality of the signal received by UEs.
- Define a statistical model that can be used to predict the level of error experienced by users and verify its accuracy using link-level simulation.
- Explore the relation of the predicted error level to the achievable DL data rates and energy consumption of the base station.
- Demonstrate that the proposed predictive model could be used to provide input for making distortion-aware scheduling decisions.
- Explore the feasibility of utilizing the power headroom available in a subset of antennas to improve the performance of a per antenna PAPR reduction approach.

Definition of these research objectives is directly connected with theses of this dissertation, which are formulated in Section 1.2.

1.2. Theses

The theses of this dissertation, related with the research objectives listed in Section 1.1, are as follows:

1. The error level of the signal received by the user depends on the structure of the base station, particularly the peak power capabilities of the power amplifiers, the number of transceivers, and the array geometry, but is mainly determined by the power allocation and precoding of the transmitted signal.
2. The balance between the achievable downlink data rates and the energy consumption of the base station can be improved by distortion-aware user scheduling decisions.

3. The distortion introduced by peak power limitation in a subset of antennas can be partially compensated by a different subset of antennas with power headroom.

The first thesis, explored in Chapter 4, refers to the part of the error (in the signal received by the UE) which results from the distortion introduced in the base station. Consideration of the error level, instead of the data rate, allows to isolate the interaction of distortion and transmission channel, without the need to deal with the effects of interference and receiver (RX) noise. This kind of analysis of a multi-user MIMO (MU-MIMO) transmission in an isolated single cell provides a foundation for evaluation of the data rate performance.

The practical impact of a given level of distortion reaching the UE can only be judged in reference to the channel quality experienced by the UE. It is the combined influence of the interference, noise, and distortion that determines the achievable DL data rate. The potential for improving the balance between the DL data rates and energy consumption of the base station by making distortion-aware user scheduling decisions is demonstrated in Chapter 5, covering the second thesis of this dissertation.

The understanding of the radiation characteristic of nonlinear distortion can be alternatively used to consider a distortion compensation method. The third thesis, explored in Chapter 6, relies on exploiting antennas with power headroom to provide means for transmitting compensation signal.

A summary of contributions, related with the outlined theses, is provided in Section 1.3 together with an overview of the structure of this dissertation.

1.3. Overview of the Structure and Contributions of the Dissertation

First, the background related with the achievable data rates and the energy consumption as key performance indicators is outlined in Chapter 2. The significance of considering the power of the received distortion, in addition to the power of the received interference and noise, as the limiting factor for the user experienced data rates is indicated there.

A critical analysis of the state-of-the-art, relevant from the perspective of the research covered by this dissertation, is provided in Chapter 3. The general background on the Massive MIMO concept provides as a foundation for considering the impact of the residual TX impairments on the in-band performance metrics, with a special attention to the effects of peak power limitation. The cognitive gaps identified in the state-of-the-art are connected with the theses of this dissertation in Section 3.4.

While each of the three theses are handled primarily in dedicated chapters (from Chapter 4 to Chapter 6), all of them focus on a limiter-based PAPR reduction as the primary source of nonlinear distortion. There is especially strong connection between Chapter 4, which is dedicated to characterizing the impact of distortion on the DL signal quality assuming ideal channel conditions for the user, and Chapter 5, which provides insights into the combined effects of interference, noise, and distortion.

Final conclusions, covering the complete body of research contained in this dissertation, are provided in Chapter 7. Additionally, each of the thesis-dedicated chapters ends with a section detailing the summary and proof of thesis (i.e., Section 4.8 for the first thesis, Section 5.4 for the second thesis, and finally Section 6.6 for the third thesis).

The first contribution of this dissertation is the proposal of statistical models which could be used for predicting the in-band performance impact of the nonlinear distortion introduced in TX of Massive MIMO base station. Definition of these models, as well as their validation in link-level simulation, is covered in Chapter 4. While the relations with the achievable data rates and energy consumption of the base station are explored in Chapter 5.

The second contribution of this dissertation, covered in Chapter 6, is the proposal of a novel method of using antennas having power headroom for compensation of the distortion introduced by limiter-based PAPR reduction. The implementation potential of the proposed approach was confirmed by the decision made by Nokia to file a patent application for protection of this solution.

1.4. List of Prepared Publications and Patent Application

The intermediate results, observations, and conclusions were presented by Łukasz Skomra (the PhD candidate) at two conferences. The related papers (co-authored with the supervisor and the assistant supervisor, with at least 70% contribution from the PhD candidate) published in conference proceedings (both with 70 points on the list of the Polish Ministry of Science and Higher Education), are:

- Ł. Skomra, B. Jelonek and K. Staniec, “On the impact of residual transmitter distortions in 5G massive MIMO base stations,” in *2023 International Wireless Communications and Mobile Computing (IWCMC)*, 2023, pp. 380–385, doi: 10.1109/IWCMC58020.2023.10182434.

This article considers the fundamental aspects of the way the distortion introduced by the residual TX impairments can be expected to impact the quality of the signal received by UEs. It shows that with increasing number of antennas on the base station

side the performance can generally be expected to improve. However, it also demonstrates that the impact on the user experienced data rates depends on the number of spatially multiplexed UEs and channel characteristics. The initial analysis captured in this article is expanded in Chapter 4 and Chapter 5 of this dissertation.

- Ł. Skomra, B. Jelonnek and K. Staniec, “Is antenna reservation superior to increasing input back-off in 5G massive MIMO base stations?,” in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023, pp. 1–6, doi: 10.1109/VTC2023-Fall60731.2023.10333415.

This article studies the performance of a PAPR reduction method that uses reserved antennas for compensating nonlinear distortion introduced by the operation of peak power limitation. The simulation results demonstrate that the compensation performance varies depending on the spatial scenario. Additionally, the article points out that usage of spatially uncorrelated channels leads to more favorable characterization of the studied solution, compared with the performance observed in spatially correlated channels. The analysis captured in this article is not directly reflected in this dissertation. However, the novel solution for compensating nonlinear distortion, proposed in Chapter 6, is to some extent similar to the compensation method studied in that paper.

Apart from these scientific articles, following patent application was filed on 18.03.2024 by Nokia for an invention proposed by the PhD candidate and the assistant supervisor (with at least 50% of contribution from the PhD candidate):

- Ł. Skomra, B. Jelonnek, “Apparatuses and methods for compensation of transmitter distortions.”

This patent application introduces the novel solution for compensating nonlinear distortion, which is described in Chapter 6.

Chapter 2

Achievable Data Rates and Energy

Consumption as Key Performance Indicators

The achievable DL data rate and the energy consumption of 5G Massive MIMO base stations are the key performance indicators that are considered in this work. In fact, the relevance of these aspects is reflected in the fact that out of eight parameters describing the key capabilities of a 5G network in the vision of the International Mobile Telecommunications for 2020 and beyond (IMT-2020), captured in [3], four are related with the UE data rates and cell throughput (i.e., peak data rate, user experienced data rate, spectrum efficiency, and area traffic capacity) and another one covers the energy efficiency (with a focus given to the energy consumption on the network side).

A general consideration of the characteristics of Massive MIMO base stations relevant from the perspective of the achievable DL user data rates and cell throughput is covered in Section 2.1, while the relation to the energy consumption of the DL processing chain is outlined in Section 2.2.

2.1. Achievable Downlink User Data Rates and Cell Throughput

The channel capacity, resulting from the Shannon's noisy channel coding theorem, defines a limit for the maximum data rate of reliable transmission [4, Ch. 5]. In cellular networks, in addition to the noise power N at the user, also the power of the received interference I (possibly both inter- and intra-cell) should be considered when calculating the achievable spectral efficiency (SE), which is limited by the channel capacity.

When the interference is treated as an additional source of noise at the RX (which is a strategy that can be used in practice in cellular networks, achieving optimal performance in the low-interference regime [1, Sec. 1.2]) the channel capacity for a deterministic additive white Gaussian noise (AWGN) channel can be described as follows:

$$C = \log_2 \left(1 + \frac{S}{I + N} \right) = \log_2(1 + \text{SINR}) \quad (2.1)$$

where C is the channel capacity (measured in bit per channel use), S is the received wanted signal power, I is the received interference power, and N is the noise power at the RX. In a cellular network, the signal-to-interference-plus-noise ratio (SINR), which is a linear power

ratio between the signal power S and the sum of the interference power I and noise power N , can be used to represent the channel quality experienced by the UE [4, Ch. 6.7.4]. When considering a single isolated cell, the channel quality can be characterized by the signal-to-noise ratio (SNR) when the intra-cell interference can be neglected.

The Massive MIMO base station precodes the TX signal to achieve a coherent combination at the UE location. This increases the SNR, compared to incoherent combination, by up to the maximum array gain (equal to the number of antennas in the array). Hence, the large number of antennas on the base station side may enable an increase in the achievable SE, by influencing the SNR portion of the channel capacity in (2.1).

In practice, the processing accuracy of the HW used in the transceiver chains of both the base station and the UE is limited. Hence, the potential impact of distortion on the achievable SE should also be considered, in addition to the dependence on noise and interference. When the distortion is treated as additional noise at the RX, similarly to the way the interference is handled in (2.1), the power of the RX distortion may be handled in conjunction with the power of the RX noise, however the power of the TX distortion may need to be generally treated separately, since the TX distortion goes through the channel. With these assumptions, the limitation of the SE for a deterministic AWGN channel can be represented by

$$SE = \log_2 \left(1 + \frac{S}{I + N + D} \right) = \log_2(1 + \text{SINDR}) \quad (2.2)$$

where D is the received distortion power. The signal-to-interference-plus-noise-plus-distortion ratio (SINDR) represents the overall quality of transmission to the UE, covering also the influence of the distortion introduced on the base station side.

Since it is the relation of the received power of the wanted signal S to the received power of distortion D that impacts the achievable SE, a characterization of the extent of a coherent combination (i.e., the effectively applicable array gain) of the distortion radiated from a Massive MIMO base station is fundamentally important. That aspect, explored in Chapter 4, influencing the signal-to-distortion ratio (SDR) part of (2.2) is at the core of this dissertation.

The massive number of antennas on the base station side not only provides the array gain that can improve the channel capacity by increasing the SNR, but it makes a high degree of spatial multiplexing feasible. On the one hand, this enables the base station to allocate all the available frequency resources to each of the UEs, without any multiplexing of UEs over frequency, which can significantly increase the user experienced data rates. On the other hand, that high degree of spatial multiplexing enables a significant increase in the cell throughput (i.e., the sum data rate of all the UEs in a cell), potentially approaching linear

scaling with the number of multiplexed UEs [1, Sec. 1.3.4] (in addition to the logarithmical scaling of the channel capacity in (2.1), achieved by increasing the received power levels).

The performance related with the achievable DL data rates (both the user experienced data rates and the cell throughput) is not only impacted by the number of UEs that can be spatially multiplexed in practice, but also by the way this spatial multiplexing impacts the signal and distortion power at each UE. The consideration of the impact of the SINR conditions on the user achievable data rates and cell throughput is covered in Chapter 5.

2.2. Base Station Energy Consumption in Downlink Transmission

Most of the electric energy consumption in mobile networks is attributed to the base station sites, with about 50% of the energy consumed by the base station equipment itself and about 40% spent by cooling [5]. With initial deployments of 5G base stations, the share of the energy consumed in the radio unit was observed to increase to about 90%, from about 80% in 4G, with about 10% of the electric energy consumed in the baseband unit [5]. The significance of the role the base station has in the overall network electricity consumption was reflected in the 3GPP Rel-18 study on network energy savings for NR [6], focused on improving operation in low and medium load cases.

With significantly increasing the number of antennas, compared with conventional base stations, the energy consumption in each of the per antenna transceiver chains as well as in the digital signal processing handling the Massive MIMO transmission needs to be kept in check. In fact, challenging goals for the balance between the achievable data rates and energy consumption were defined for the future extreme MIMO systems in [7]. With an order of magnitude higher peak capacity envisioned to be available with less than twice the energy consumption of the state-of-the-art 5G Massive MIMO systems.

Naturally, it can be expected for the load in different cells of the network to vary during both the daily and weekly cycles, due to different patterns of user activity, mostly resulting from the characteristics of the human-induced network traffic. While a considerable attention is given by the industry to supporting mobile network operators in the efforts to reduce energy consumption in the low and medium load cases, this dissertation focuses on the high load regime. Such operating mode may occur naturally during peak hours in highly loaded cells. Additionally, a high instantaneous load may be induced during low and medium load periods by energy saving techniques like micro-discontinuous transmission [7].

In the high load operation, the relatively high PAPR of an OFDM signal (which is normally higher than 10 dB) is problematic from the perspective of the energy efficiency of

the DL transmission. Using an input back-off (IBO) equivalent to the PAPR seems to be a waste of resources, both from the perspective of the energy consumption and the cost of the HW components used in a Massive MIMO base station. This comes from the fact that increase in the IBO not only reduces the energy efficiency of the transmission chain, but also results in using components with increased peak power capability (in order to provide enough mean output power to attain the required cell range). Therefore, the research effort covered by this dissertation focuses on the PAPR reduction (done in alignment with the peak power capabilities of the PAs), which enables energy-efficient operation of the base station.

Chapter 3

State-of-the-Art

This dissertation is anchored in the field of wireless communication in cellular networks. The state-of-the-art relevant from the point of view of the undertaken research covers the aspects related with the significance of the residual TX impairments, from the perspective of the DL transmission from a Massive MIMO base station to UEs. Consideration of the impact of nonlinear distortion introduced by a limiter-based PAPR reduction is the primary focal point.

The general background, presented in Section 3.1, establishes the context for consideration of the way the relevance of the residual TX impairments has been viewed by the Massive MIMO research community and how the understanding of the impact of the HW non-idealities was evolving, which is explained in Section 3.2. That part of the state-of-the-art provides foundation for the research objectives handled in Chapter 4 and Chapter 5 of this dissertation.

Exploration of the spatial degrees of freedom (available in Massive MIMO transmission) to approach the problem of PAPR reduction, covered in Section 3.3, provides basis for the research objective covered in Chapter 6.

The cognitive gaps, identified in the state-of-the-art, are connected with the theses of this dissertation in Section 3.4.

3.1. General Background

The concept of a noncooperative cellular multi-user MIMO system introduced by T. Marzetta in 2010 [8], referred in the later literature as Massive MIMO [9], is widely considered as pivotal for increasing the SE and reducing the energy consumption of cellular networks [10]. The primary characteristic of such a system is the usage of base stations with large number of antennas to spatially multiplex a large (but smaller than the number of antennas) number of UEs. A Massive MIMO network was envisioned in [8] to operate in a time-division duplex (TDD) mode, utilizing uplink (UL) pilots to acquire the channel state information (CSI) which can be used by simple linear receive combining and transmit precoding algorithms. Authors in [10] highlighted the scalability of such a system, compared with a classical MU-MIMO where the numbers of base station antennas and UEs are in the

same order of magnitude. An overview of the history of development of MU-MIMO, leading to the Massive MIMO concept was presented in [11].

While the canonical Massive MIMO system [12] is imagined to operate in a TDD mode, with base stations transmitting to single-antenna UEs, it is not excluded to use FDD (with own challenges related with the lack of reciprocity between the forward and reverse channels) and multi-antenna UEs (by virtually combining multiple single-antenna UEs into multi-antenna UE [13]), as already hinted in [8].

While the analysis in [8] dealt with an unlimited number of antennas and a fixed number of UEs, it was later clarified in [10] that the Massive MIMO concept is envisioned to utilize few hundred of antennas transmitting to tens of UEs (utilizing the same time-frequency resources), per cell. This suggests that the number of UEs is expected to be roughly an order of magnitude smaller than the number of antennas. However, it was demonstrated in the example in [12] that an optimum sum SE in an independent and identically distributed (i.i.d.) Rayleigh fading channel was achieved, for UEs in equal SNR conditions of -5 dB, with a ratio of antennas to UEs of about 5 for zero forcing (ZF) and about 2 for maximum ratio (MR) processing, with a relatively wide range of ratios good for MR (from 1 to 5) but a narrower range suitable for ZF (particularly due to a reduced channel null space when the ratio of 1 was approached). This suggests that, potentially depending on the SNR conditions, the optimally supported number of UEs may be closer to the number of antennas than an order of magnitude.

Authors in [14] demonstrated that when channels are spatially correlated, which was an assumption that was different than used in a majority of earlier research contributions basing on spatially uncorrelated channels, the seemingly fundamental limitation of SE caused by the pilot contamination in a multi-cell network, highlighted in [8], can be eliminated by employing multi-cell signal processing (i.e., receive combining and transmit precoding) based not only on intra-cell but also inter-cell channel estimates. The significance of these results is emphasized by the fact that practical channels are expected to be spatially correlated, as demonstrated in multiple measurement-based studies referenced in [13].

A definition of a Massive MIMO cellular network given in [13], motivated by the initial commercial deployments, describes a network consisting of at least 2 cells operating in TDD mode, utilizing base stations with at least 64 antennas, employing linear combining and precoding schemes capable of spatially multiplexing at least 8 UEs per cell. Authors in [13], highlighted the importance of evolving the Massive MIMO concept, to “Massive MIMO 2.0”, by considering spatial channel correlation and using multi-cell signal processing

(capable of suppressing both intra- and inter-channel interference). While the Massive MIMO base stations used in the initial 5G deployments employ arrays of 32 or 64 antennas [11], with dedicated transceivers allowing for a fully digital beamforming, the utilized rectangular array geometry (e.g., with 8 columns and 4 rows of cross-polarized antennas [11]) leads to arguably coarse spatial resolution over azimuth angles, compared to horizontally oriented uniform linear arrays (which are widely studied in the Massive MIMO literature), with a similar number of antennas.

With the “Massive MIMO 2.0” concept, outlined in [13], the evolution of Massive MIMO networks using compact arrays of co-located antennas is considered to be in the hands of industry, supported by academia [11], while the primary research effort is foreseen to be shifting to future disruptive areas. Some of the potential research directions are highlighted in [13], with the focus on the large intelligent surfaces (both active and passive – intelligent reflecting surfaces), post-cellular network architectures (cell-free networks), and sub-THz communications. These are in line with a part of the future research topics listed in [11]. There, extremely large aperture arrays (characterized by hundreds of distributed antennas, covering cell-free Massive MIMO as a special case) are considered, where most of the communication is expected to happen in the radiating near-field, instead of the far-field experienced with the conventionally deployed compact Massive MIMO arrays of co-located antennas. Another group of research topics is represented by Holographic Massive MIMO (characterized by antenna surface with approximately continuous aperture, covering the large intelligent surfaces as a special case). Use cases for positioning, radar, and using machine learning to optimize communication in Massive MIMO networks are also mentioned in [11] as relevant development directions.

3.2. Residual Transmitter Impairments in Massive MIMO Base Stations

While the assumption of ideal HW simplifies the network-level analysis, it may hide the aspects relevant from the perspective of the achievable performance and lead to design decisions that are suboptimal from the perspective of practical devices with non-ideal HW. In the papers considering the impact of the HW non-idealities, the conclusions on the significance of the residual TX impairments in Massive MIMO base stations evolved from assuming uncorrelated noise, vanishing with increasing the number of antennas, to embracing that nonlinear distortion radiate to distinct directions and in the worst case can achieve the same array gain as the wanted data signal.

Assuming ideal HW.

When the HW non-idealities are ignored, as in the seminal paper [8] and some other key papers (e.g., [15] and [16]), their potentially limiting impact cannot be observed. In that case, signal-to-noise-ratios (SNRs) of all UEs improve, without limit, when the number of antennas increase while the sum output power is not changed. Hence, the inter-cell interference may be identified as the only remaining limitation for increasing SE, provided that the intra-cell interference is handled by precoding. However, when the HW impairments are considered, their impact depends on how they scale with the array gain.

Expecting decreasing impact of the TX impairments with increasing array gain.

According to [10], the per antenna constraints on the processing accuracy and linearity of HW could be relaxed, since the related negative effects average out when the over-the-air performance is considered. The correctness of this conclusion depends on the validity of the assumption that the errors introduced in different antennas do not combine coherently at the RX. While the phase noise was identified in [10] as a remaining limiting factor, which could be potentially dealt with at the RX side, its impact was not expected to be more problematic than in legacy systems, as concluded in [12].

In a key paper [17], a new system model for Massive MIMO transmission was considered. It covers the effects of the residual HW impairments in both the TX, similarly to the model studied in [18] for conventional point-to-point MIMO systems, and the RX. Motivated by earlier experimental studies, the combined effect of different transceiver impairments (assumed to be partially corrected for by compensation algorithms) were modeled in [17] as additive Gaussian noise, independent of the wanted signal but proportional to its power per antenna. While there may be types of distortion that are not well represented by such a simplified model, the comparison between stochastic and deterministic models in [19] seems to confirm that the additive Gaussian noise model provides adequate accuracy. The fundamental conclusion in [17] is that the residual HW impairments limit the achievable data rates in Massive MIMO systems, contrary to the claims in [8] where the asymptotical analysis using infinite number of impairment-free antennas led to a conclusion that the only remaining relevant limitation for increasing SE is the pilot contamination.

While the residual HW impairments in both base station and UE impact the SE, proportionally to the level of distortion introduced at both sides of the link, the authors in [17] claim that only the residual impairments at the UE side are limiting the achievable data rates, when the number of antennas at the base station side grows large. This is motivated by

the underlying assumption, reflected in modelling distortion using uncorrelated Gaussian noise, that the errors introduced in different antennas do not combine coherently. It follows, that the quality requirements for the HW components used in the base station could be relaxed, proportionally to the number of antennas.

The possibility to allow for increasing the level of tolerable HW impairments in Massive MIMO base stations, in comparison to conventional MIMO systems, was reiterated in [19]. That study, apart from considering distortion and noise introduced per antenna by PAs and digital-to-analog converters (DACs), which were identified by the authors as the two key HW impairments, additionally explored the effects of mutual coupling between antennas. For a deterministic behavioral model covering the PA and the effects of mutual coupling, about 6 dB improvement of the error vector magnitude (EVM) with increasing the array gain by 6 dB could be observed in the results provided for 4 UEs in an i.i.d. Rayleigh channel. While the results for a Rician channel, from the same behavioral model, for the case of a half-wavelength spacing between the array elements are somewhat harder to interpret (since the angular separation between UEs seems to be considerably smaller than the half-power beamwidth), an EVM improvement both with increasing the number of antennas and with reducing the power in the LoS paths could be observed. The conclusion that the relation between the EVM and the array gain for the combined contribution of the studied deterministic models (covering PA, mutual coupling, and DAC) seems to be well approximated by the stochastic models (both additive and multiplicative) was backed-up by simulation results for 4 UEs (in Rayleigh and LoS channels) with a surprisingly high level of residual EVM, which can be observed in [19] for the results with no HW impairments.

Clarifying that the residual TX impairments may combine coherently at the user.

An analysis based on deterministic behavioral models of PAs in [20] lead to a conclusion that the nonlinear distortion can generally combine coherently over-the-air, which was not exposed in the observations of the impact of PA nonlinearities in [19], marking a pivot in the perception of the impact of the residual TX impairments in the Massive MIMO research community. This was backed up by the theoretical analysis in [21] which revealed that nonlinear distortion from Massive MIMO base station is generally expected to be radiated into distinct directions. This was demonstrated already earlier for active antenna arrays, based on the analysis of the radiation patterns of the intermodulation products, in [22] (in the context of satellite communications) and in [23] (covering broad-band arrays with a wide field-of-view). The extension of the two-tone analysis from [21] with a two-carrier I/Q modulated signals in [24], exposed that the out-of-band (OOB) emissions in case of multi-

UE transmission can be expected to be also beamformed to the same directions as the intended UEs. In fact, it is the direction of users where these emissions are always the strongest according to [24].

The worst-case scenario, identified in [21], was a transmission to a single user over a single-path LoS channel. In this case, both the wanted signal and nonlinear distortion are expected to experience the same array gain in the UE direction, which is a fundamentally different behavior than anticipated based on the relatively widely established model of uncorrelated noise from [17]. Although it is the multi-user transmission that is associated with Massive MIMO by default, the results obtained for a single-user case in LoS conditions may be highly relevant in practice, like explained in [25], when the transmit power is distributed between UEs based on their path losses to achieve comparable quality of service (QoS).

The analysis in [25], based on simulation results from a PA model (fitted to measured data), demonstrates that with a relatively small number of beams in the precoded TX signal there are certain unfavorable directions (at the UEs and other distinct angles) with a considerable power OOB, while in most of the directions that power is relatively low. When the number of beams was increased, the directivity of the OOB radiation pattern was reduced, resulting in lack of distinct unfavorable directions, approaching the radiation pattern of a single array element. The qualitative observations from [25] were confirmed and extended by the quantitative analysis in [2], covering not only OOB, but also in-band power of distortions.

A method to predict the output of antenna arrays when isolators between PAs and antennas are removed was introduced in [26] and extended in [27]. The distortion resulting from mutual coupling between array elements and impedance mismatches between PAs and antennas was demonstrated to have non-negligible impact [26], compared to the distortion levels of single PAs, especially when the in-band part of the spectrum was considered [27]. The relevancy of covering the effects related with impedance mismatch and mutual coupling, when isolators are not used, was also strongly highlighted in [28], where authors suggested that a “well-designed” Massive MIMO system can tolerate higher levels of nonlinear distortion per antenna, compared to a non-beamforming system.

Expecting the array gain of distortion to reduce when increasing the number of radiation directions with significant power.

A comprehensive study of the spatial characteristics of distortion, described in [2], demonstrated that the radiation pattern of distortion (from a nonlinear PA, modeled using a

memory polynomial), both in-band and OOB, can vary between omni-directional and fully aligned with the wanted signal pattern. Fundamentally, considering the directions in which the different classes of intermodulation products introduced in PAs are expected to be radiated, based on the earlier analysis for active arrays in [22] and [23], authors in [2] used precoding weights to calculate the radiation directions of distortion. The relation of the number of radiation directions with significant power in the precoded signal to the number of antennas was identified in [2] as the primary factor influencing the variation of the power of distortion over the steering angles. This means that for frequency-selective channels the nonlinear distortion could be expected to be radiated with less directivity than for frequency-flat channels, as noted already for OOB radiation in [25]. In fact, it was established in [2], that the number of radiation directions of distortion grows as a cube of the number of UEs and a square of the number of channel taps (with significant power), which leads to a uniform radiation pattern when the number of directions is larger than the number of antennas.

The spatial distribution of OOB distortion was studied in [29], based on simulation using PA model as the source of nonlinearity, for transmission in line-of-sight (LoS) and non-line-of-sight (NLoS) conditions using realistic channel models, specifically targeting for identification of intermodulation (IM) beams. In the case of a single user free-space LoS scenario no IM beams were observed, confirming the conclusions of the theoretical analysis from [21]. While it was demonstrated that the coherent combining of distortion reduced when the number of directions of the wanted signal was increased, a uniform distribution of OOB distortion in space was not achieved, for the uniform linear array (ULA) with 128 antennas used in the verification, even with more than 10 UEs in NLoS channels. Authors in [29] suggest that distortion-aware user scheduling and precoding could be useful to guarantee fulfilment of regulatory limits, considering that the emission levels strongly depend on the use case and position at which the emissions are measured. Similar effects of nonlinear distortion were identified for the distributed Massive MIMO, however more uniform distribution of distortion power in space (compared with a co-located deployment) was observed [30]. Contrary to co-located arrays, the distortion from distributed arrays was demonstrated in [30] to combine coherently in specific zones instead of certain directions. However, for a transmission to 1 UE in a free-space LoS channel a full coherent combination of distortion at the UE position can be expected also in the distributed Massive MIMO [30].

The analytical framework introduced in [31] for exploration of the behavior of nonlinear distortion, covering single-user transmission in channels with uncorrelated Rayleigh fading, confirmed that an increase in the number of transmission paths leads to a reduction in the

power of received distortion, which is in line with the observation in [2] of more uniformly radiating distortion when the number of taps with significant power in the precoding filter is increased. Another conclusion in [31], based on both analytical calculations and numerical simulation results, was that when the delay spread reduces the power of the received in-band distortion increases, which is aligned with the observations in [2] that the directivity of distortion decreases with increasing delay spread. It was highlighted in [31] that the level of received OOB distortion reduces faster than the in-band distortion, both with increasing the number of paths as well as with decreasing the maximum delay spread (which was demonstrated in [31] for the case of 20 multipath components).

An analytical method for identification of the directions of radiation of nonlinear distortion (introduced by PAs) from active reconfigurable intelligent surfaces (RISs) was proposed in [32]. It was extended in [33] (not published in a peer reviewed journal at the time of writing this dissertation, as of August 2024), to cover extremely large aperture arrays (ELLAs) and operation in the radiative near-field (extending the model with a depth domain), in addition to the far-field operation considered initially in [32]. Based on a comparison of analytical results (covering third-order nonlinearities) against numerical simulations, it was shown in [33] that the angular directions of UEs in azimuth do not influence the elevation angles of distortion, while the azimuth angles of distortion are impacted by the elevation positions of UEs. Moreover, for the radiative near-field it was shown that the focal points of distortion are at different distances from the base station than the distances at which UEs are positioned.

Using the knowledge about the way the residual TX impairments radiate to either optimize scheduling decision or influence their radiation patterns.

While [2] hinted at the possibility to use prediction of the way the residual TX impairments radiate as an input for a distortion-aware scheduling, no details were given, apart from stating that UEs with similar channels should be avoided to be assigned SCs close to each other. However, this kind of solution was described in more detail by the authors of [2] in a patent application [34].

It was also suggested in [2] that the distortion could be steered away from the users, at the cost of reducing the array gain of the wanted signal, without exploring this aspect further. The knowledge of the spatial characteristics of the third-order impairments, assumed to be acquired during transmission, was used in [35] to perform a distortion-aware precoding. The utilized iterative approach for calculation of the precoding weights was demonstrated, in numerical simulation for NLoS scenarios, to lead to a better sum rate compared to both MR

and ZF precoding, at the cost of increased complexity. The improvement in the sum rate, demonstrated in [35], clearly depends on the number of UEs, which is likely related with the degrees of freedom available for spatial distribution of distortion away from UEs. For example, for the highest presented SNR of 20 dB, the sum rate improvement against ZF is about 51% for 2 UEs, while it is about 24% for 4 UEs.

A family of zero third-order distortion (Z3RO) precoders, introduced in [36] and covered in more detail in [37], is an alternative solution for reducing the level of distortions radiated toward UEs, using a closed-form expression to calculate the linear precoding weights. The Z3RO approach aims to achieve cancellation of the distortion at UE locations using a set of saturated antennas to provide a destructive superposition of distortion at the UE, based on a simplified expectation of third-order IMs, without using actual PA model and the statistics of the actual wanted signal. This comes at the cost of reduced array gain and increased radiation of distortion to directions other than occupied by the UEs, which may be significant in power in case only a single antenna is used for compensation. The simulation results presented in [37], covering only a single UE scenarios, demonstrate that Z3RO was able to perform better than MR precoding in a free-space LoS channel only when the operating point for the non-saturated antennas was relatively close to the peak power limit (about 5 dB or less), while it performed worse for larger back-off values (i.e., at lower power operating points). While for NLoS channel, the cross-over point between Z3RO and MR performance was demonstrated in [37] to be achieved for a wider range of power operating points (already at about 8 dB back-off when compared to MR with linearized PA), it may be an optimistic result due to usage of i.i.d. Rayleigh channel. The simulation results covered in [38], obtained using measurement-based channel models, demonstrate that the reduction of the level of distortion received by UEs, achieved with Z3RO compared with MR precoding (without PA linearization), was smaller for 2 UEs than for 1 UE, while using two times more saturated antennas (4 instead of 2, out of a total of 32). This suggests that increasing the number of UEs may lead to a limited performance improvement, which seem to be in line with observation that could be made based on the results shared in [35] for a different scheme of distortion-aware precoding.

3.3. Peak-to-Average-Power Ratio Reduction in Massive MIMO

For classical non-beamforming base stations, there is a plethora of different techniques that could be utilized to reduce the PAPR of the DL signal, in order to improve the energy efficiency working point of the PA in the TX chain. Traditional (operating per antenna)

solutions addressing the issue of a high PAPR of an OFDM signal, were grouped into three distinct classes in [39]: signal distortion, multiple signaling & probabilistic, and coding. While it was concluded in [39] that there was no single approach that could be considered as the best from all the perspectives covered in that analysis (i.e., bit error rate increase, bit rate loss, implementation complexity, and power increase), the technique of clipping and filtering (CF), falling into the signal distortion category, was singled-out as a low-complexity solution providing a good PAPR reduction performance for OFDM systems. Although the traditional PAPR reduction techniques operating per antenna could be successfully utilized in Massive MIMO base stations, the excess degrees of freedom available in a Massive MIMO channel (resulting from about an order of magnitude more antennas on the base station side in relation to the number of spatially multiplexed UEs) could be exploited to substantially reduce the PAPR of the per antenna signal (achieving a constant envelope in an extreme case) [10].

Different classes of solutions can be identified when focusing on the PAPR reduction methods that do not require any special handling on the UE side. Categories of techniques based on the following primary distinguishing features are considered here: joint precoding and PAPR reduction; PAPR-reducing perturbation vectors; signal distortion with a channel-null constraint; tone reservation; antenna reservation.

Joint precoding and PAPR reduction, as an optimization problem.

The generation of a precoded time-domain per antenna signal with a constrained PAPR (or peak power) based on a per UE frequency-domain data input was proposed in [40] to be treated as an optimization problem, leading to a PAPR-aware precoding. The non-convex problem of the PAPR minimization was reformulated in [40] as a convex optimization problem of minimization of the magnitude of the time-domain signal, covering precoding (including handling the inter-user interference), OFDM modulation, and PAPR reduction. Since the computational complexity of the fast iterative truncation algorithm (FITRA) proposed in [40] to tackle a relaxed version of that optimization problem is widely considered to be excessively high from a practical deployment perspective [41], there is a number of other approaches for the joint precoding and PAPR reduction (JPP) proposed in the literature, focused on improving the complexity while solving the same (or an equivalent) optimization problem. Notably, a remarkable complexity reduction of more than 97% in relation to FITRA, while maintaining a comparable PAPR reduction and symbol error rate (SER) performance, was reported for the accelerated proximal gradient algorithm (APGM) proposed in [42].

Recently, the linearized alternating direction method of multipliers (LADMM) was proposed in [43] to directly address the non-convex PAPR minimization problem of the PAPR-aware precoding, without the relaxations taken to tackle the convex problem of the peak power minimization in [40]. The computational complexity of LADMM for reaching PAPR of 4 dB was demonstrated in [43] to be about 2% of FITRA [40] and about 82% of APGM [42], while achieving comparable PAPR reduction and SER performance, including the performance with imperfect CSI.

The common characteristic of the JPP approaches targeting the PAPR reduction without negatively impacting the EVM and OOB emissions is the increase in power of the precoded signal in relation to ZF precoding solution. With a constraint on the sum output power, this has a practical consequence of allocating less power to the data signal, resulting in a degradation of the SINR, as pointed out in [40] and explored in more detail in [41]. While the approach taken in [41] was proven to deliver a minimum PAPR solution, the analysis using a novel problem formulation and method providing control over the trade-off between the PAPR reduction and the related power increase proposed in [44] revealed that, in fact, an optimal balance between the PAPR reduction and the power increase can be achieved by methods taking minimization of the peak power as an optimization target.

PAPR reduction using optimized perturbation vectors.

While most of the optimization-based solutions to the PAPR reduction problem rely solely on the precoding operation to constrain the PAPR of the time-domain signal [40] – [44], there are also alternative approaches that separately consider the problem of optimization of the PAPR-reducing signals.

A perturbation-assisted scheme for a PAPR reduction independent from the precoding design proposed in [45] relies on adding optimized perturbation symbols, constrained to the null space of the MIMO channel matrix, to the precoded frequency-domain data symbols. Based on simulation results with ZF precoding, it was shown that the PROXINF-ADMM algorithm proposed in [45] required an order of magnitude less iterations (200 instead of 2000) than FITRA [40] to achieve a similar PAPR reduction and bit error rate (BER) performance. More importantly, a similar improvement was demonstrated for the run time (measured as a proxy for the computational complexity) required for reaching a given PAPR level. In fact, for moderate levels of PAPR reduction, a substantially faster convergence of PROXINF-ADMM compared to FITRA was clearly demonstrated in [45], with only several iterations required by PROXINF-ADMM to reach a PAPR of 6 dB (a level that was reached after about 350 iterations by FITRA). Additionally, the inaccuracy of the CSI was shown in

[45] to have a negligible influence on the PAPR reduction performance, as expected, while the BER performance degradation for PROXINF-ADMM was on a par with the performance achieved by FITRA.

Despite the fact that the perturbation-assisted JPP approach proposed in [46] allows for an optional separation of precoding and PAPR reduction, in line with the strategy followed in [45], its fundamental advantage leading to a reduced computational complexity is claimed to be the joint optimization of the precoded data signals and the PAPR reduction signals. The JPP, formulated as a convex optimization problem, was solved in [46] by the proposed MU-PP-GDm algorithm (based on the gradient descent with momentum) which entails both a generation of the frequency-domain precoded signal and an optimization of the frequency-domain PAPR-reducing perturbation signals (restricted to the null space of the MIMO channel) using the same criterion as in the ZF precoding. The computational complexity required to reach a PAPR of about 4 dB by MU-PP-GDm was reported to be only about 0.1% of FITRA [40] and about 3.1% of PROXINF-ADMM [45]. A distinguishing feature of MU-PP-GDm, resulting from the way the JPP problem was formulated in [46], is the lack of a noticeable increase in the power of the precoded signal compared to the ZF precoding.

PAPR reduction based on the signal distortion with a channel-null constraint.

The class of solutions to the optimization problems formulated for generation of perturbation vectors in both [45] and [46] maps to iterative clipping with the PAPR-reducing signal constrained to the used tones (avoiding the OOB emissions) within the null space of the Massive MIMO channel (keeping the goal of limiting the inter-user interference), according to the mentioned references. Actually, there is a group of techniques utilizing directly the conventional per antenna PAPR reduction approaches based on signal distortion in combination with projecting the PAPR-reducing signals to the null space of the channel.

The performance of the CF with a channel-null constraint (CFCNC) was verified in Massive MIMO use cases (having about an order of magnitude more antennas than UEs) in [47], based on an algorithm consisting of iteratively applied: beamforming, CF, and projection of the PAPR reduction signal to the null space of the MIMO channel. It was concluded that the proposed method was proven to be especially effective when the number of antennas was large, the required PAPR was low, and the SNR was high. However, based on the shared simulation results it can be observed that the negative impact of the PAPR reduction on the average capacity was visibly improved for the CFCNC compared with the traditional per antenna CF [39] only in a subset of the covered range of the average PAPR (from about 5 dB up to about 9 dB, with the largest benefit visible for about 6 dB PAPR).

The impact of the variation of the mean power over antennas on the performance of a peak-limiter based PAPR reduction, not considered in [47], was studied for the CFCNC in [48] where a novel beamforming method with a per antenna power constraint (PAPC) was realized using a firefly algorithm (FA), reducing the variation the mean power between antennas, was proposed. It was demonstrated in [48] that the proposed combination of the FA-based beamforming and the CFCNC improved the relation between the achievable average throughput and the average PAPR when the required PAPR was lower than 7 dB, but led to a slight degradation of throughput compared to the ZF precoding for higher PAPR levels. Interestingly, based on the shared simulation results, it seems that actually the traditional per antenna CF solution [39] benefited significantly more than the CFCNC from the combination with the FA-based beamforming, which is clearly visible in the apparently reduced throughput gap between the CF and the CFCNC. However, it is worth to point out that the usage of a frequency-flat Rayleigh fading channel might have influenced the results in a way that favored beamforming with PAPC over the regular ZF precoding, as it could be expected based on the results from [2] that the power variation over antennas increases when the channel delay spread is reduced.

In order to reduce the complexity of a limiter based PAPR reduction with a channel-null constraint, a PCCNC approach based on an adaptive peak cancellation (PC) [49], instead of the CF used in [47], was proposed in [50]. While a visibly lower computational complexity of PCCNC compared to CFCNC was demonstrated in [50], both the PAPR reduction and throughput performance (in a wide range of PAPRs) were on similar levels. Two modifications of the original PCCNC [50] were proposed in [51] to further reduce the complexity and improve the convergence rate of this PC-based PAPR reduction solution: division of the array into blocks of antennas which are separately parallelly processed with PCCNC; handling multiple peaks of the time-domain signal in a single iteration of PCCNC. While both increase in the number of antenna blocks, effectively reducing the size of the channel matrix that needs to be processed to enforce the channel-null constraint, and increase in the number of peaks handled at a time were demonstrated in [51] to lead to a noticeably lower complexity, both led to a minor degradation in the BER performance. The comparison of the complexity required for reaching a given PAPR level with 100 antennas for PCCNC using 5 blocks to the optimization-based MU-PP-GDM [46] revealed that the PCCNC needed less real multiplications for down to about 5.2 dB PAPR in case of 1 peak at a time handled by PCCNC iteration, while for lower PAPR levels MU-PP-GDM outperformed PCCNC. For the multi-peak approach for PCCNC the convergence was visibly improved,

leading to a lower complexity than MU-PP-GDM down to a level about 0.1 dB above the PAPR reduction threshold of 4 dB when 4 peaks were handled at a time. It is worth to note that the PAPR reduction performance in frequency-selective channels, shared in [51], seems to be negatively impacted by using many relatively small blocks of antennas, based on a cross-over between the PAPR reduction performance with 10 blocks of 10 antennas and the original PCCNC using a single block, which is not observable in the simulation results for frequency-flat channels published in an earlier conference paper [52]. Interestingly, a noticeable degradation of BER performance for MU-PP-GDM in relation to ZF precoding is visible in the simulation results reported in [51], which is not that apparent based on the SER performance shared in [46] (for a range of SNRs between 0 and 10 dB, instead of 20 to 35 dB covered in [51]).

Based on the observation that the level of correlation between the PC signal vector (spanning over all antennas at the time instant of a PAPR reduction event) before projection to the null space of the MIMO channel and the precoded data signal vector should be low in order to achieve a noticeable PAPR reduction, a modification of the PCCNC method from [50] was explored in [53]. A novel method for the PC signal vector generation proposed in [53] achieves that reduced correlation by increasing the instantaneous power in antennas with power below a configurable threshold at the time instances of the PAPR reduction event occurrence in any of the other antennas. While a significantly faster convergence of the proposed method was demonstrated, the throughput results shared in [53] seem to paint a mixed picture, especially when a relatively high threshold for the power increase of 4 dB was used. However, for the average PAPR of about 6.5 dB and lower, the throughput results for the novel method of the PC signal generation proposed in [53] demonstrate a consistent improvement over the original PCCNC approach from [50].

Tone reservation.

A conventional per antenna tone reservation (TR) technique [39] could be successfully used to reduce the PAPR in Massive MIMO base station, provided that the fundamental downside of the capacity reduction resulting from less tones available for allocation to data signals in a TR-based PAPR reduction approach is acceptable. One way to exploit TR in combination with the additional degrees of freedom offered by the Massive MIMO channel could be to relax the optimization problem of JPP, by allowing for the reserved tones to be used for the PAPR reduction, as pointed out in [40]. In fact, a TR-like approach proposed in [54] addressed the problem of minimization of the PAPR of the time-domain signal (using

dummy symbols injected into the tones that are empty in at least a subset of MIMO layers) combined with minimization of the power reserve utilized by the PAPR-reducing signal.

A relatively straightforward take on reducing the computational complexity of a TR-based solution in a multi-antenna use case was made with the adaptive tone reservation (ATR) proposed in [55], where the iterative reduction of the PAPR was performed only on the antenna with the highest PAPR in a given iteration, instead of all the antennas. It is worth to point out that the demonstrated improvement in the convergence speed of the ATR over the conventional per antenna TR [39], resulting from the significant variation of the PAPR over antennas observed in simulation in [55], may be noticeably impacted by the channel characteristic and the method used to calculate the precoding weights.

Antenna reservation.

An alternative strategy for making use of the extra spatial degrees of freedom, proposed in [56], hinges on reserving a subset of antennas for transmitting solely compensation signals for the distortion introduced by the other set of antennas that transmit the data signals. While in this case the expected data rate related performance improvement is to some extent limited by the need to compensate for the reduction of the array gain applicable to the data signals, it is expected that this kind of solution may entail a relatively low computational complexity due to no need for an iterative handling of the spatial degrees of freedom (used in other solution, either for generating a PAPR-reduced signal directly or projecting the PAPR-reducing signal to the channel null space). According to [56], that reduction in the array gain leads to a less severe negative effect on the spectral efficiency than observable for TR-based approaches. The verification of the performance of the antenna reservation (AR) compensating the distortion introduced by a per antenna limiter-based PAPR reduction was expanded in [57] by considering the iterative clipping and filtering, in addition to the hard clipping covered in [56].

An extension of the AR-based compensation [56], to cover not only a different per antenna PAPR reduction approach (based on companding transforms) but also distortion introduced by nonlinear PAs, was proposed in [58], where also the feasibility of using the proposed AR-based approach with imperfect CSI was demonstrated. Usage of AR-based approach for compensation of distortion introduced by a per antenna PC-based PAPR reduction was covered in [59], where a clear improvement in both the PAPR reduction and the BER performance in relation to the per antenna PC-based PAPR reduction without AR [49] was demonstrated. An optimal level of performance in terms of potentially achievable data rates was obtained in [59] with about 20–25% of reserved antennas, comparable to the

observation in [58], with less antennas needed in case of only a partial compensation based on different EVM requirements in different MIMO layers. It was demonstrated in [60] that not only the number of antennas but also the selection of antennas for reservation, from all the antennas in a uniform planar array, mattered when a spatially correlated MIMO channel was assumed. While an improvement in the BER performance was achieved, by optimizing which of the antennas were dedicated for transmitting the compensation signals, for scenarios with UEs at equal angular spacing, there was no visible impact observed for randomly distributed UEs.

3.4. Cognitive Gaps and Associated Theses

The research covered by this dissertation revolves around consideration of the practical impact the distortion introduced in the TX chains of a Massive MIMO base station have on the key performance indicators related with the DL data rates and the network energy consumption. Interestingly, a number of recent papers on the user-centric cell-free Massive MIMO [61] go back to the approach to model the impacts of the residual impairments based on uncorrelated Gaussian noise, same as used in [17], which was proven already in [20] as incorrectly describing the effects of nonlinear distortion introduced by PAs, especially for LoS channels and relatively small number of UEs [2]. It could be argued that when fully linearized PAs in combination with distortion-free PAPR reduction solutions are used, the model based on uncorrelated noise may be accurate enough. However, the assumption of a distortion-free operation leads to over-constraining, since in practice UEs can be expected to tolerate some level of distortion with negligible impact on the achievable data rates, due to limited SINRs.

The importance of considering spatially correlated channels was highlighted in [13], while the practical relevance of LoS scenarios was argued in [12]. However, in fact, the vast majority of studies proposing novel solutions for PAPR reduction based on exploiting the excess degrees of freedom of the Massive MIMO channel, covered in Section 3.3, used spatially uncorrelated Rayleigh channels (frequency-selective in most of the cases) to demonstrate the performance of the proposed solutions, with a notable exception of [60]. The behavior of nonlinear distortion was primarily explored so far using PA models, as explained in Section 3.2, with the attention focused on the problem of the OOB emissions.

The first gap identified in the current state-of-the-art is the need for characterization of the EVM performance of a conventional limiter-based per antenna PAPR reduction solution in Massive MIMO scenarios with spatially correlated channels. It appears to be especially

the case for channels with a significant LoS component and for transmission from planar arrays. The aspect of the limited angular granularity, resulting from the array geometry (especially with the not particularly massive number of columns and rows in the antenna arrays used in the initial commercial Massive MIMO deployments [11]), seems to require attention from the EVM performance perspective.

The second identified gap is consideration of the impact that the power allocation to UEs have on the way the residual TX impairments influence the EVM performance. While the vast majority of studies covered in Section 3.2 used an equal power allocation strategy, [2] and [25] were the only identified studies that practically scratched the surface of the problem of an unequal power distribution, by providing general observation that the nonlinear distortion in that case radiate primarily to the users with higher power allocation.

These two gaps are addressed first from the EVM performance perspective in Chapter 4, in relation with the first thesis of this dissertation (which aims at clarifying the fundamental behavior of the distortion introduced in the TX of Massive MIMO base station). Next, the DL data rate performance perspective is explored in Chapter 5, in relation to the second thesis of this dissertation (which points to the potential limitation of the achievable DL data rates resulting from the TX distortion, impacting the balance between the data rates and the energy consumption of the base station).

The third gap identified in the current state-of-the-art, related with the second thesis of this dissertation addressed in Chapter 5, is exploitation of the opportunity to utilize the knowledge of the radiation characteristic of the nonlinear distortion to improve the spatial-temporal scheduling. This expands on the opportunity of a distortion-aware frequency scheduling, hinted in [2] and followed up in [34].

The final gap identified in the current state-of-the-art, related with the third thesis of this dissertation addressed in Chapter 6, is the potential opportunity of exploiting the power headroom in a subset of antennas to provide compensation for distortion introduced in a different subset of antennas. Instead of modifying precoding to reduce the variation of power over antennas, as in the FA-based beamforming proposed in [48], the idea is to utilize the power variation to improve the EVM performance of a conventional per antenna PAPR reduction, using a similar mechanism to AR-based solutions [56]–[60] but without dedicating any antennas for transmission of the compensation signals.

Chapter 4

Predicting the In-Band Impact of the Residual Transmitter Impairments

In practical base stations, the non-idealities of the HW components lead to the TX signal impairments, which alter the signals that are intended to be transmitted from antennas of a base station. Usually, the negative effects of impairments are partially corrected by digital signal processing on the TX side. The uncompensated (residual) HW impairments, combined with the non-idealities of the digital signal processing, introduce distortion into the data signal. In-band, the impact of distortion is visible as an error in the transmitted signal. While OOB, it constitutes unwanted emissions, which must be limited due to spectral emission requirements related with coexistence (between different channels or different radio access technologies).

This chapter addresses the first thesis of the dissertation. First, it aims to clarify whether the distortion radiated from Massive MIMO base stations can restrict the achievable DL data rates. Then, it considers the impacts of the structure of the base station (i.e., peak power capabilities of PAs, number of antennas, and the array geometry) and selected characteristics of the transmitted signal (i.e., power level and precoding).

Two defined statistical models provide predictions of the level of EVM resulting from nonlinear distortion. Their prediction accuracy is validated in link-level simulation. One of the models reflects in a simplified way the impact of the array geometry, by considering the number of antennas, while the other one implicitly covers the dependence on the array geometry, by basing on the channel and precoding matrices. The insights obtained based on the EVM performance in link-level simulations form a foundation for extending the analysis to the system-level, handled in Chapter 5. A characterization of the EVM performance with quantization noise is also made.

The basis for this research is the original work of the PhD candidate published in [62] (co-authored with the supervisor and the assistant supervisor). While the vast majority of the simulation results presented in Section 4.7 were not published before, some of the results from [62] were reused, as explicitly indicated in the text.

4.1. System Model

The model considered here describes a system that represents a single cellular base station transmitting to multiple UEs located in one cell. The base station uses an array of M antennas to transmit the DL signal to K single-antenna UEs.

The signals received by K UEs are linear combinations of the signals transmitted from the base station on all M antennas, passed through a MIMO channel, as visualized in Figure 4.1 for transmission on one of J subcarriers.

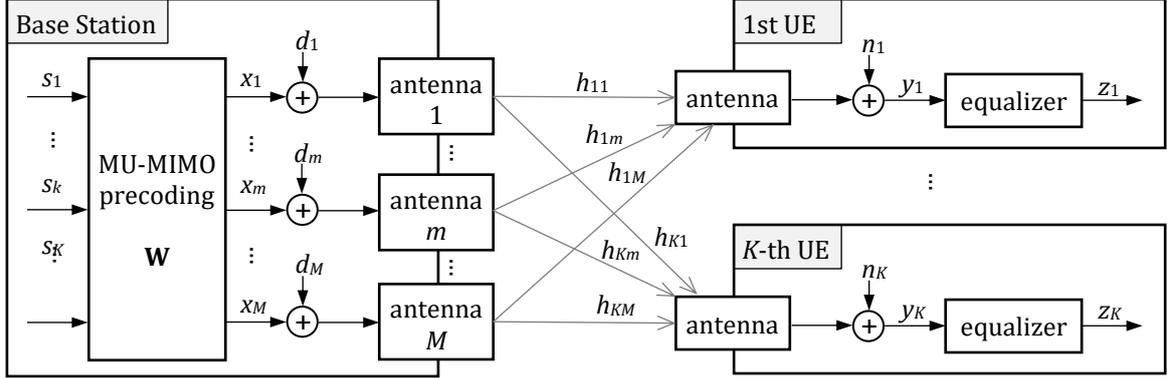


Figure 4.1: Base station, using an array of M antennas, transmitting distorted signals to K single-antenna users.

The frequency-domain data samples s_k are transmitted, according to the 5G New Radio (NR) physical layer [63], within resource elements (REs) organized in orthogonal frequency-division multiplexing (OFDM) subcarriers (SCs) in frequency and OFDM symbols in time.

For a given OFDM symbol, the frequency-domain signals received by K UEs on the j -th SC can be described as follows:

$$\mathbf{y} = \mathbf{H}(\mathbf{x} + \mathbf{d}) + \mathbf{n} \quad (4.1)$$

where \mathbf{y} is a $K \times 1$ vector of frequency-domain data samples received by K UEs, \mathbf{H} is a $K \times M$ matrix of channel responses from M base station's antennas to K UEs, \mathbf{x} is an $M \times 1$ vector of the precoded data samples, \mathbf{d} is an $M \times 1$ vector representing TX distortion, and \mathbf{n} is a $K \times 1$ vector of the RX noise. Note that the subscript related with indexing over SCs is omitted, for brevity, when the interaction between SCs does not directly matter.

Each UE receives a signal over a multiple-input single-output (MISO) channel. The frequency-domain signal y that is received by the k -th UE on j -th SC for a given OFDM symbol, can be described as follows:

$$y_k = \mathbf{h}_k^T (\mathbf{x} + \mathbf{d}) + n_k \quad (4.2)$$

where $y_k \in \mathbb{C}$ is a frequency-domain data sample received by the k -th UE, $\mathbf{h}_k^T \in \mathbb{C}^{1 \times M}$ represents channel responses from M base station's antennas to the UE antenna, $\mathbf{x} \in \mathbb{C}^{M \times 1}$

is a vector of the precoded data samples, $\mathbf{d} \in \mathbb{C}^{M \times 1}$ is a vector representing TX distortion, and $n_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_k^2)$ represents the RX noise (with σ_k^2 being the noise power). The received signal y_k is equalized by the k -th UE using pilot symbols, transmitted along with the data symbols, which results in signal z_k .

For each of the J SCs, the precoded data vector \mathbf{x} is constructed by summing up the K per UE data samples s_k multiplied by the K per UE precoding vectors $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$:

$$\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k \quad (4.3)$$

Therefore, the signal received by the k -th UE contains the wanted signal part (carrying the data samples for the k -th UE, received via the effective scalar channel $g_k = \mathbf{h}_k^T \mathbf{w}_k$), the intra-cell interference, the received TX distortion, and the RX noise:

$$y_k = \underbrace{\mathbf{h}_k^T \mathbf{w}_k s_k}_{\text{Wanted signal}} + \underbrace{\sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{h}_k^T \mathbf{w}_i s_i}_{\text{Intra-cell interference}} + \underbrace{\mathbf{h}_k^T \mathbf{d}}_{\text{Transmitter distortion}} + \underbrace{n_k}_{\text{Receiver noise}} \quad (4.4)$$

Here, the power allocation to UEs and SCs is performed using the precoding weights. While the data samples, selected from the symbols of a quadrature amplitude modulation (QAM) constellation (with the constellation size of either 4, 16, 64, 256, or 1024 [63, Sec. 7.3.1.2], depending on the modulation order selected for transmission in a given RE), are normalized such that $\mathbb{E}\{|s_k|^2\} = 1$.

The explanation of the considered power allocation strategies in Section 4.1.1 is followed by a description of the utilized models of the residual TX impairments, provided in Section 4.2.

4.1.1. Allocation of Power and Frequency Resources

Clearly, the distribution of the available TX power between UEs impacts their SINRs. At the same time, it influences the relation between the powers of the received wanted signal and distortion [62]. Hence, the power allocation strategy is an important aspect of the system model when the impact of residual TX impairments is considered.

Here, the sum power constraint (SPC) level of P is the same irrespective of the number of antennas M (i.e., operation in a high-power mode is assumed [25]). This means that with more antennas in the array the mean power over antennas is reduced, therefore PAs used in the TX chains can have lower output power capabilities, but the equivalent isotropic radiated power (EIRP) is increased due to increased directivity. An alternative would be to consider

a low-power mode for the array [25], in which the SPC is reduced with increasing the number of antennas (enabling reduction of the energy consumption), but the EIRP is not increased.

The mean output power per antenna is P/M . Precoding weights distribute the power between antennas, without explicit mean per antenna power constraint (PAPC). This generally leads to variation of the mean power between antennas, as visible in Figure 4.2. Which means that the energy efficiency operating points are generally different on different antennas and are changing when the precoding weights change, which can happen each slot.

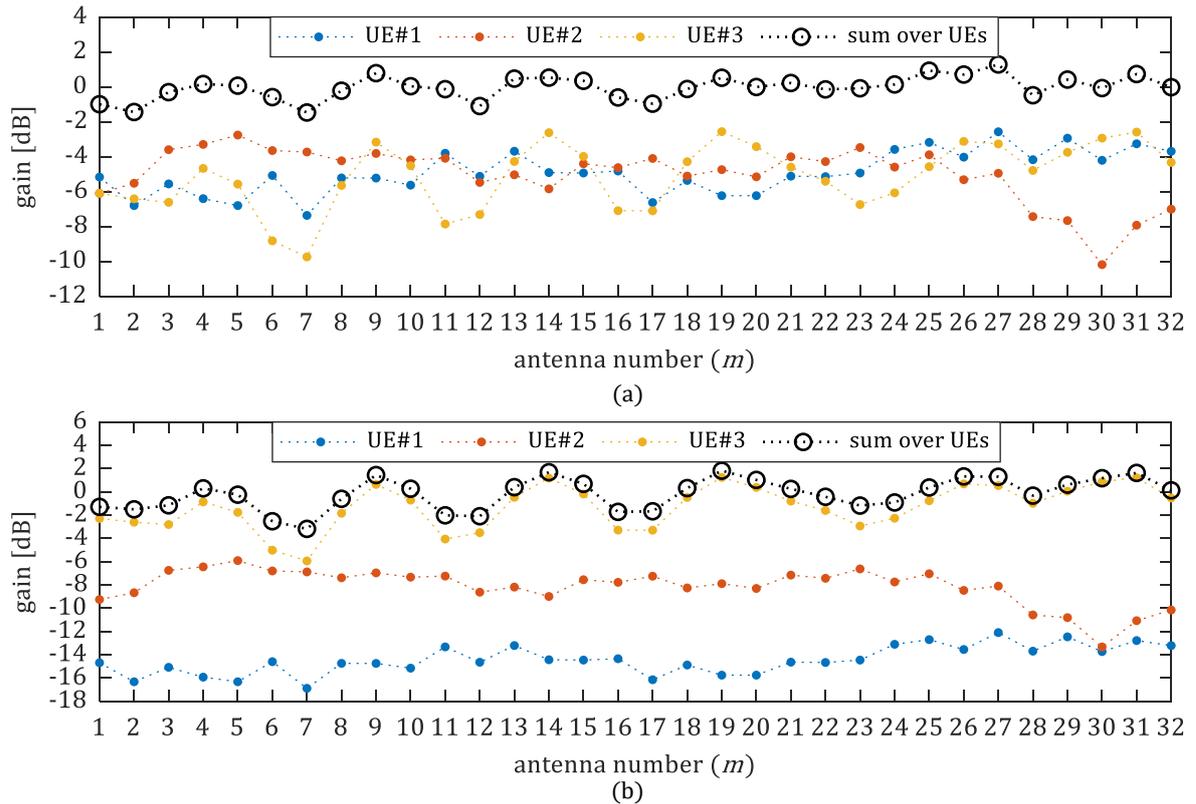


Figure 4.2: Normalized gain of the per antenna precoding weights for 3 UEs in multi-path LoS channels: (a) equal per UE power; (b) power allocated in a way to compensate the differences between path losses.

The per UE precoding vectors \mathbf{w}_k distribute the power P (available for transmission of an OFDM symbol), between K users (i.e., transmission layers) over M antennas for each of the J SCs (i.e., frequency resources). Each UE, scheduled in a given slot in time (covering 14 OFDM symbols), is allocated all the available SCs. Hence there is no multiplexing of UEs in the frequency domain, which is in line with the approach for allocation of resources in Massive MIMO networks, as envisioned in [10].

The power is allocated per UE in a way that the full available output power P is utilized:

$$\sum_{k=1}^K p_k = P \quad (4.5)$$

where p_k is the per UE power budget.

Two different approaches for power distribution between UEs are considered:

- 1) Equal distribution of power between UEs, as exemplified in Figure 4.2(a) where the mean gain over antennas is $10\log_{10}(1/K) = -4.8$ dB for each of the UEs.
- 2) Distributed in a way to achieve a comparable QoS (by allocating power inversely proportional to SNRs obtained with equal power distribution), as visualized with the example in Figure 4.2(b) for UEs located 100, 200, and 400 m away from the base station, having mean gains over antennas of -14.3 , -7.9 , and -1.0 dB, respectively.

The strategy of equal power distribution between UEs is assumed by default, with p_k equal to P/K . This leads to keeping the differences between the gains of the per UE effective scalar channels, resulting from the differences in the per UE path loss. The alternative option (where UEs are allocated different power budgets, depending on their path loss conditions), can lead to equal achievable SE, when the impact of TX distortion is neglected.

The mean output power over SCs is P/J . Two different strategies of the power allocation between SCs that are considered here:

- 1) For each UE, the channel gains over SCs are equalized. This results in variation of TX power over SCs, exemplified in Figure 4.3(a) for a UE power budget of P/K .
- 2) Each SC is allocated the same power budget P/J , which leads to a flat TX power over SCs, as visualized in Figure 4.3(b) for equal power distribution between UEs.

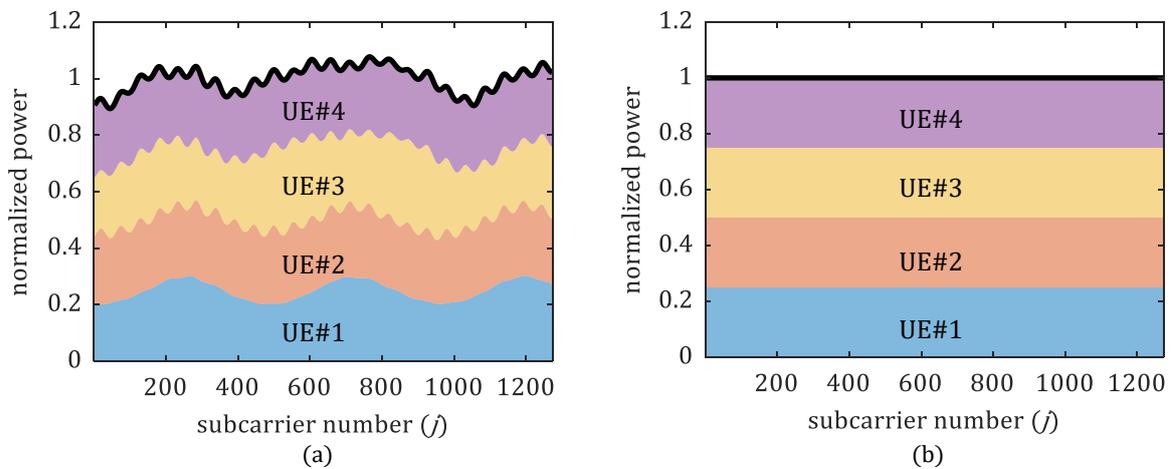


Figure 4.3: Power distribution between J subcarriers and K users: (a) example of an equal power distribution between users, equalizing channel gains; (b) equal power distribution between both users and subcarriers.

4.2. Models of the Residual Transmitter Impairments

Two distinct classes of residual TX impairments models are considered in this study. Behavioral models of peak power limitation (resulting from either saturation in the PA or an explicit PAPR reduction), covered in Section 4.2.1, capture the effects of nonlinear distortion. A contrasting view is provided using a model based on additive Gaussian noise to reflect the effects of quantization, as explained in Section 4.2.2.

All of the considered models of residual TX impairments operate per antenna, on the time-domain representation of the precoded signal x_m which is obtained after the CP-OFDM modulation.

4.2.1. Nonlinear Distortion Introduced by Peak Limitation

The PAs used in the TX chains have certain output power capabilities. The peak output power is limited by the saturation level, while the mean output power can be constrained by the thermal design of the base station. Since a saturation in the PA leads to OOB emissions, a controlled limitation of the peak power is preferred. This can be achieved by the iterative clipping and filtering (ICF), which is a class of PAPR reduction methods based on signal distortion [39].

The clipping operation is modeled, based on the definition in [39], as:

$$\text{clip}(x_m[n]) = \begin{cases} x_m[n], & |x_m[n]| \leq \text{CL} \\ \text{CL} e^{j\angle x_m[n]}, & |x_m[n]| > \text{CL} \end{cases} \quad (4.6)$$

where $x_m[n]$ is a time-domain sample on the m -th TX path, CL is the amplitude clipping level, and $\angle x_m[n]$ is the angle of $x_m[n]$. In ICF, clipping followed by filtering (i.e., bandwidth limitation) is applied multiple times, as visualized in the diagram in Figure 4.4.

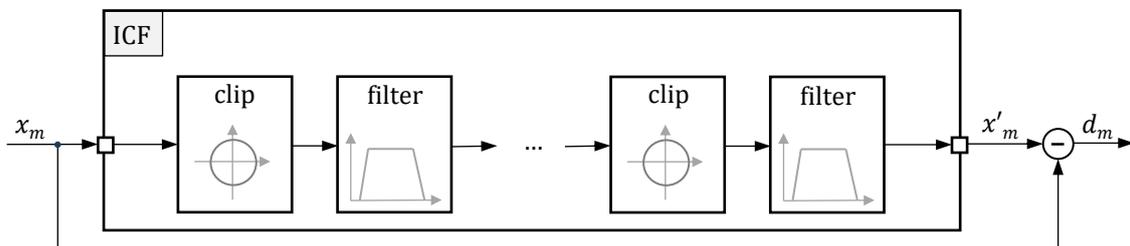


Figure 4.4: Block diagram of iterative clipping and filtering (ICF).

The hard clipping (HC) operation in (4.6), not followed by filtering, is used here to model the effects of saturation in a linearized PA. A perfect linearization is assumed, like in [37]. In that case the CL is selected in alignment with the peak power capability of the PA.

While the input back-off (IBO) is selected in a way that the expected level of the input power maps to the nominal average output power $P_{\text{nom}} = P/M$, as visualized in Figure 4.5.

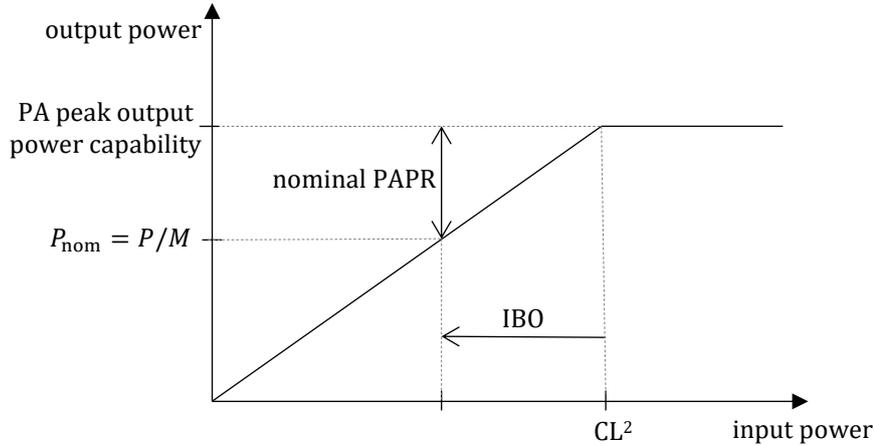


Figure 4.5: Clipping level (CL) and input back-off (IBO) in relation to average and peak output powers of a linearized power amplifier (PA) and the nominal peak-to-average power ratio (PAPR).

The limitation of the instantaneous power, performed in the clipping block of both HC and ICF, results in a reduction of the mean power level. Here, this power loss is not corrected for. This is a deviation compared to [1], where it is assumed that the compensation of the HW impairments is done in a way that the mean output power is not modified.

The ratio between the CL^2 and the expected level of the input power per antenna is called the clipping ratio (CR) here. It is usually expressed in decibel, as follows:

$$CR = 10 \log_{10} \left(\frac{CL^2}{\mathbb{E}\{\|x_m\|^2\}} \right). \quad (4.7)$$

4.2.2. Noise Added by Quantization

The quantization error can be introduced into the wanted signal in the digital part of the DL processing chain, up to the DAC. Here, the focus is given to the per antenna quantization of the time-domain OFDM signal, hence the effects of the potential quantization in the frequency-domain part of the processing chain (i.e., before the CP-OFDM) are not considered.

It is assumed that quantization is performed in a way that the correlation of the quantization errors between antennas is avoided, for example, by using dithering [64, Sec. 13.8.2]. The per antenna quantization noise is modeled here as an additive complex Gaussian, with a zero mean and a variance resulting from the expected quantization noise power. This kind of model is utilized to focus on the combined influence of different

quantization points in the DL processing chain, instead of considering the relation of the quantization noise power with the bitwidths and sample rates.

4.3. Preliminaries

Before introducing the proposed EVM performance prediction models, covered in Section 4.4, the background related with the tolerable EVM, beam steering, radiation patterns, and intermodulation is outlined in the following subsections.

4.3.1. Tolerable Error Vector Magnitude

The EVM is a metric that can be used to quantify the negative effects of TX distortion on the received signal quality. The expected level of EVM can be defined as

$$\text{EVM} = \sqrt{\frac{\mathbb{E}\{|z - s|^2\}}{\mathbb{E}\{|s|^2\}}} \quad (4.8)$$

where s represents the transmitted modulation symbols and z represents the corresponding received and equalized modulation symbols. Since here $\mathbb{E}\{|s|^2\} = 1$, as explained in Section 4.1, the calculation of the expected EVM can be simplified to

$$\text{EVM} = \sqrt{\mathbb{E}\{|z - s|^2\}}. \quad (4.9)$$

The requirements on the modulation quality for the base station TX, measured using EVM, are defined by 3GPP in relation to the modulation order [65]: 17.5% for QPSK, 12.5% for 16QAM, 8.0% for 64QAM, 3.5% for 256QAM, and 2.5% for 1024QAM. The conformance with these requirements is verified in a controlled environment of an anechoic chamber. The 3GPP allows a test system uncertainty of $\pm 1\%$ for the modulation quality measurement (relaxing the acceptance level of EVM by 1 percentage point) [66], hence it may be challenging to accurately measure the impact of the TX distortion when considering levels of about -40 dBc (i.e., 1%) or lower.

In practice, the error vector ($e = z - s$) represents the combined influence of interference, RX noise, and TX distortion. Therefore, in order to focus the analysis on TX distortion, a controlled environment is needed in which the influence of the interference and RX noise, represented jointly by SINR, can be separated from TX distortion. This is why usage of simulation models, in which isolation of different sources of error is feasible, can be advantageous at the stage of exploratory research.

For a given level of SNR required for successful transmission, depending on the UE conditions represented by SINR, there is some level of EVM that can be tolerated, as

demonstrated in Figure 4.6. On the one hand, this means that the TX does not need to be distortion-free. This was clear already when looking at the modulation order specific requirements for EVM defined by 3GPP. On the other hand, the level of tolerable EVM depends on the headroom between the actual SINR and the required SNR, which is influenced by the selection of the modulation order and code rate for transmission. And the smaller this headroom is the closer the tolerable EVM is to 0%, which is not immediately apparent when considering solely the modulation order specific EVM limits defined by 3GPP. This clearly leads to a conclusion that the tolerable EVM may be lower (i.e., stricter) than required by 3GPP, depending on the SINR condition of the user.

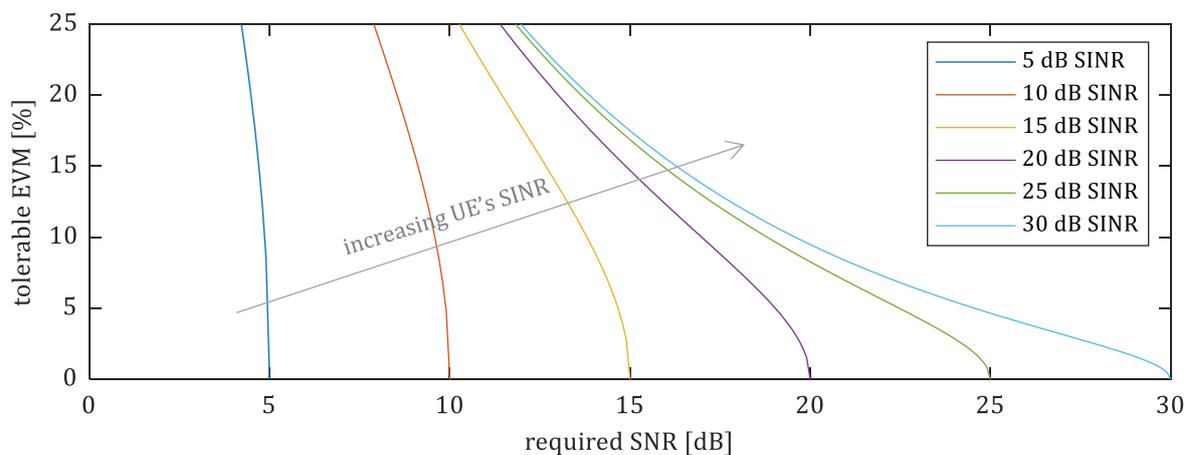


Figure 4.6: Tolerable EVM from distortion reduces when the headroom between the UE SINR and the required SNR decreases.

While it is not explicitly demonstrated here, the tolerable EVM may be relaxed compared to the 3GPP required level when a sufficiently low code rate is used in combination with a given modulation order.

4.3.2. Beam Steering and Array Factor

For an antenna array with a uniform half-wavelength spacing between antennas, positioned in the vertical y - z plane with positive values of the azimuth angle $\phi \in [-\pi, \pi]$ going from the x -axis in the direction of the y -axis and positive values of the elevation angle $\theta \in [-\pi/2, \pi/2]$ going from the horizontal x - y plane in the direction of the z -axis, beam steering (which can be considered as a special case of beamforming) requires application of phase gradients for horizontal $\Delta\gamma_y$ and vertical $\Delta\gamma_z$ elements of the array, as follows:

$$\Delta\gamma_y = -\pi \sin \phi \cos \theta \quad (4.10)$$

$$\Delta\gamma_z = -\pi \sin \theta. \quad (4.11)$$

The detailed procedure used to define (4.10) and (4.11) is covered in Appendix Section A.2.

In a practical Massive MIMO base station, the feasible steering range is narrower (i.e., it does not cover the full sphere around the base station). For a 3-sector cell-site, each individual antenna array may have the azimuth steering limited to about $\phi \in [-60^\circ, 60^\circ]$, to reduce the inter-sector interference. While the utilized elevation steering range may primarily depend on the deployment (e.g., steering mainly below horizon in a suburban setting or covering also tall buildings in an urban canyon situation) and on the mechanical tilt of the array.

Consequently, the azimuth and elevation steering angles (ϕ and θ , respectively) could be identified based on the linear phase gradients ($\Delta\gamma_y$ and $\Delta\gamma_z$) as

$$\theta = \arcsin\left(-\frac{\Delta\gamma_z + n2\pi}{\pi}\right) \quad (4.12)$$

$$\phi = \arcsin\left(-\frac{\Delta\gamma_y + n2\pi}{\pi\cos\theta}\right) \quad (4.13)$$

where $n = (0, \pm 1, \pm 2, \dots)$. The calculation of (4.12) and (4.13) is explained in detail in Appendix Section A.2.3.

For an array of identical antennas, the radiation pattern of the array is a product of the radiation pattern of a single antenna element and the array factor (AF), which describes the radiation pattern of an array of isotropic sources, as explained in more detail in Appendix A.

While a coherent combining of the signal transmitted from an array using beam steering is expected in the direction of the steering angles, as exemplified in Figure 4.7 with the main lobe pointing to azimuth $\phi = 30^\circ$ and elevation $\theta = -15^\circ$ for a square uniform rectangular array (URA) using 1024 antennas, there are also other angular directions with a partially coherent combining, resulting in multiple side lobes visible in Figure 4.7.

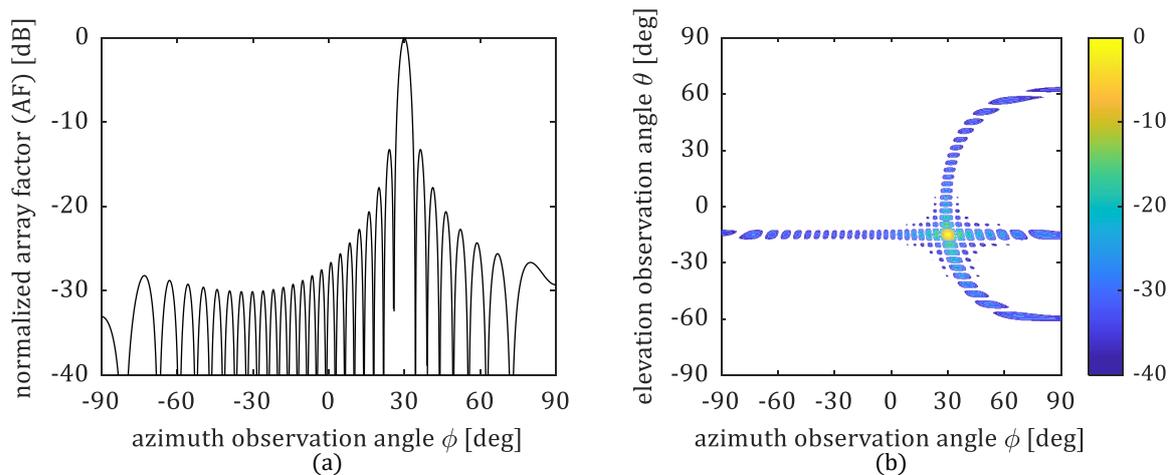


Figure 4.7: Normalized array factor (AF) for a uniform rectangular array (URA) with 32 rows and columns, steered to $\phi = 30^\circ$ and $\theta = -15^\circ$: (a) AF azimuth cut at $\theta = -15^\circ$; (b) AF, excluding values below -40 dB.

These side lobes may become relevant with precoding that does not handle inter-user interference (IUI), like MR precoding, especially when beams have unequal power allocation.

For an array with no elevation steering (i.e., $\Delta\gamma_z = 0$) and a mechanical boresight at $\theta = 0$, along the x -axis, the calculation of the horizontal phase gradient, $\Delta\gamma_y = \Delta\gamma$, and the azimuth steering angle ϕ simplifies to

$$\Delta\gamma = -\pi \sin \phi \quad (4.14)$$

$$\phi = \arcsin\left(-\frac{\Delta\gamma + n2\pi}{\pi}\right). \quad (4.15)$$

Equations for calculating phase gradients and steering angles for arrays with different than half-wavelength spacings between antenna elements are provided in Appendix A.

4.3.3. Radiated Intermodulation Products

The nonlinear distortion introduced per antenna to the TX signal radiates from the array, as explained in more detail in Appendix B. The characteristic of that radiation depends on the relations of the precoding weights [2], applied to the wanted signal to control the way the wanted signal radiates to UEs.

In case of beam steering for a horizontally oriented ULA, there is a certain linear phase gradient $\Delta\gamma$ that can be applied to adjacent antennas to steer the beam to azimuth angle ϕ , as explained in Section 4.3.2. In this situation the relations between the phase gradients applicable to different UEs determine the radiation angles of distortion [23].

The antenna array filters spatially the intermodulation products (IMPs) generated by the nonlinearities in TX chains [22], leading to directional radiation pattern of distortion. Part of the IMPs is explicitly beamformed into the same directions as the wanted data signal, while the other spatially filtered portion radiates into distinct directions that are generally different than the directions of the wanted signal, as explained in more detail in Appendix Section B.2. This is visible in Figure 4.8(a), where a distinct beam of distortion d is radiated at about 57° , when the wanted signal x is radiated to UEs placed at -30° and 10° azimuth angles.

However, it turns out that there are also spatial scenarios with multiple UEs in which such distinct IM beams are not synthesized and the complete distortion is radiated towards UEs, as visualized in Figure 4.8(b), which seems not to have been identified in earlier literature. The explanation of such worst-case spatial scenarios is tackled in Section 4.4.

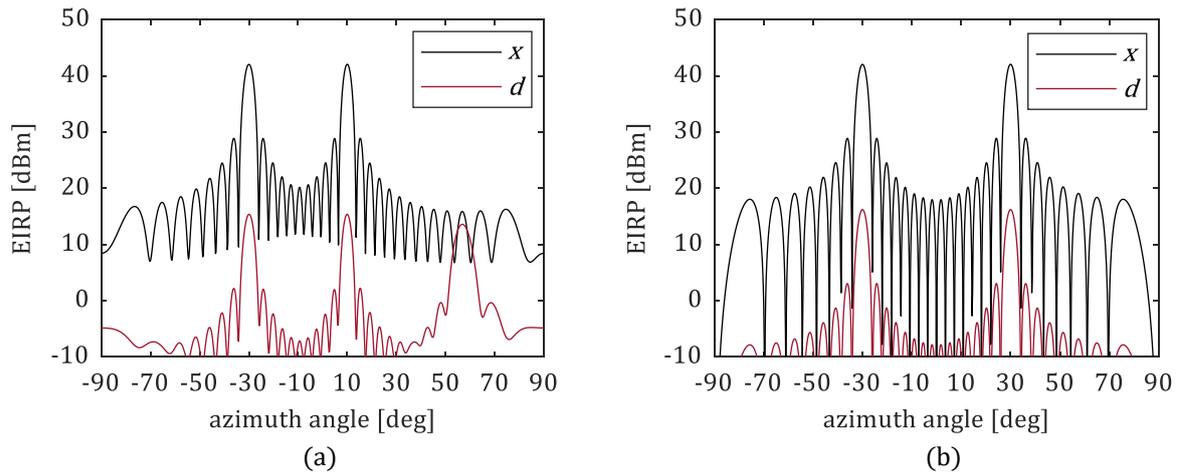


Figure 4.8: Radiation patterns of the wanted signal x and clipping distortion d for MU-MIMO transmission over free-space LoS channels to 2 UEs located at azimuth angles of: a) -30° and 10° ; b) -30° and 30° .

In case of beam steering, the ratio U of the number of the user-directed IMPs (radiating in the same directions as the wanted signal, as explained in more detail in Appendix Section B.2) to all the IMPs depends on the number of synthesized beams, which maps here to the number of UEs K . For the third-order IMPs, U can be approximated (assuming high number of SCs, used normally in OFDM) following [22] as:

$$U = \frac{(2K - 1)}{K^2}. \quad (4.16)$$

Clearly, U in (4.16) reduces when K increases. With more than 3 UEs, most of the IMPs are not explicitly spatially filtered into UE directions, as demonstrated in Table 4.1.

Table 4.1: The ratio U , of the user-directed IMPs to all the 3rd-order IMPs, from (4.16).

Number of UEs (K)	1	2	3	4	5	6	7	8
U [dB]	0.0	-1.2	-2.6	-3.6	-4.4	-5.1	-5.8	-6.3

The directivity of distortion for beamforming (i.e., going beyond the special case of beam steering) was characterized in [2], where it was pointed out that in multi-path channels the number of distinct directions to which IMPs radiate depends not only on the number of UEs, but generally on the number of radiation directions with significant power. This means on the one hand side that with low number of UEs there still can be multiple directions in case we consider NLoS. However, at the same time, if one of the UEs have LoS component and/or power is allocated to UEs such that the number of radiation directions with significant power is reduced then also the power of spatially filtered IMPs gets smaller compared to the power of the user-directed IMPs.

For beam steering, where linear phase gradients are used, the IMPs are expected to radiate at the directions defined by combinations of phase gradients of the wanted signal, as explained in more detail in Appendix Section B.2. The radiation directions of the IM beams may be calculated for horizontally-steered ULA based on (4.15), while for the general case of URA (including ULA with combined horizontal and vertical steering) based on (4.12) and (4.13).

4.4. Proposed Error Vector Magnitude Performance Prediction Models for Nonlinear Distortion in Line-of-Sight Transmission

The distortion received by the UE has negative impact on the achievable DL data rate when the EVM is higher than tolerable for the modulation order and code rate used in transmission, as explained in Section 4.3.1. In order to estimate the EVM resulting from a class of nonlinear distortion sources, which can be considered to be the most relevant for energy-efficient Massive MIMO bases stations, predictive statistical models covering transmission scenarios with LoS are proposed here. These models reflect the expected in-band performance impact of peak power limitation, resulting from both saturation in the linearized PA (modeled with HC operation) as well as a limiter-based PAPR reduction (using ICF as a method of peak power limitation).

First, a model for the per antenna EVM is considered in Section 4.4.1. It uses a Gaussian signal to predict the relation between the peak power limit and the EVM for the 5G NR transmission, based on the power allocation reflected in the precoding weights.

The result obtained from the per antenna model provides an input for the per user EVM models. The simplistic best-case prediction model, based on the number of spatially multiplexed UEs and number of antennas, provides an initial approximation of the EVM performance. That model is introduced in Section 4.4.2.1. A more elaborate model for predicting the EVM performance, based on the precoding weights, is covered in Section 4.4.2.4.

The best-case model may not be able to provide an acceptable prediction accuracy in all the cases. However, it is not excluded that it could be reasonably accurate when combined with additional rules addressing the unfavorable sets of UE angular directions, covered in Section 4.4.2.2, and the expected performance outliers, identified in Section 4.4.2.3.

While the proposed models are defined here based on free-space LoS, it is assumed that they should be able to relatively well reflect the performance in multi-path transmission with

LoS. The simulation results used to judge the prediction accuracy, for both of these channel types, are provided in Section 4.7.

It is assumed that prediction models based solely on third-order IMPs can provide acceptable accuracy of the EVM estimate for nonlinear distortion resulting from peak limitation, despite higher-order IMPs introduced in this case. That assumption is verified in simulation, against behavioral models of PAPR reduction and linearized PA, in Section 4.7.

4.4.1. Per Antenna Error

The time-domain OFDM signal can be approximated by a Gaussian signal, when the number of SCs J is high. According to [67], it is enough to use 100 SCs to have negligible approximation errors. Therefore, it is a valid approach for 5G NR carriers, which cover between 132 and 3276 SCs depending on the used channel bandwidth and spacing between SCs [65, Sec. 5.3.2].

The magnitude of a circularly symmetric complex Gaussian signal is well described by the Rayleigh distribution. This is exemplified in Figure 4.9, where a relatively good alignment between the theoretical complementary cumulative distribution function (CCDF) for Rayleigh distribution and the CCDF of the power of a time-domain signal (based on 1000 slots of randomly generated data samples for NR20, sampled at 122.88 Msps) is visible.

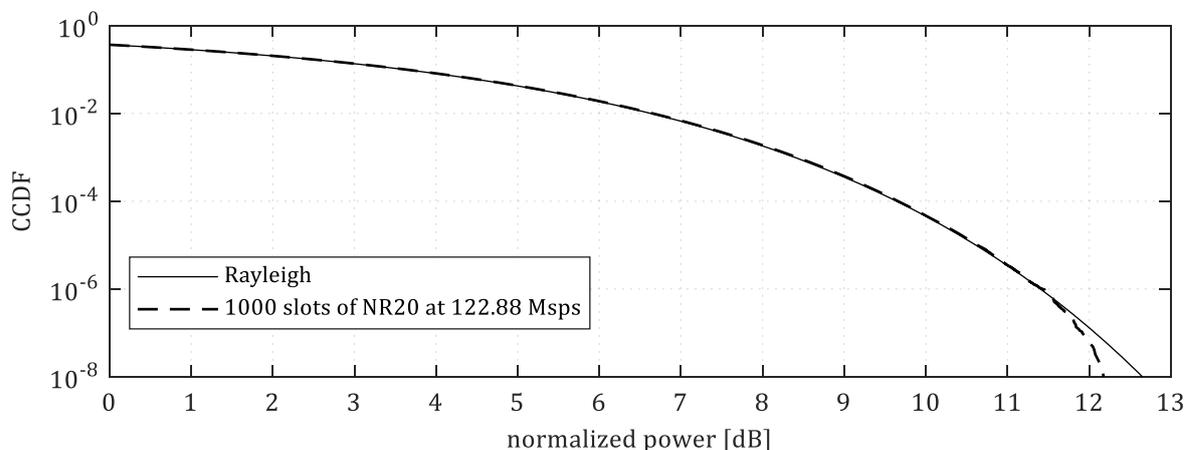


Figure 4.9: CCDF of generated NR20 data closely follows the theoretical Rayleigh distributed magnitudes.

Note that the CCDF for NR20, presented in Figure 4.9, was calculated from the complete 1000 slot long signal. When each slot is considered separately, the tail of the distribution of power values may noticeably deviate from the theoretical Rayleigh distribution. This may be relevant in practice, since the base granularity of scheduling resources to UEs in 5G NR is a single slot. However, the intention of the proposed models is to reflect the expected EVM performance, representative of a multi-slot transmission, which can be used to judge

the system-level effects of distortion on the SE performance statistics. Hence, consideration of the impact of the EVM variation between slots is not in the scope of this study.

Two types of limitation of the peak power of the time-domain TX signal are considered:

1. Saturation at the peak power capability of a linearized PA, modeled here with HC as explained in Section 4.2.1.
2. PAPR reduction using ICF, introduced in Section 4.2.1, with the peak power limit defined in relation with the peak power capability of the PA.

It is assumed that the expected level of EVM introduced by the considered sources of distortion in 5G NR carrier transmission could be sufficiently accurately predicted, by a parametric model, using the power of error introduced in a Gaussian signal when saturating it at a level corresponding with a given CL. Validity of this assumption is verified later, in Section 4.7.1.

For estimation of the per antenna EVM, the clipping operation from (4.6) was applied to $1e9$ samples of a circularly symmetric complex Gaussian signal. The relative power of error, calculated as the difference between the output and input of the clipping block (corrected with the reduction of the mean power, caused by clipping), is visualized in Figure 4.10 for a range of CRs.

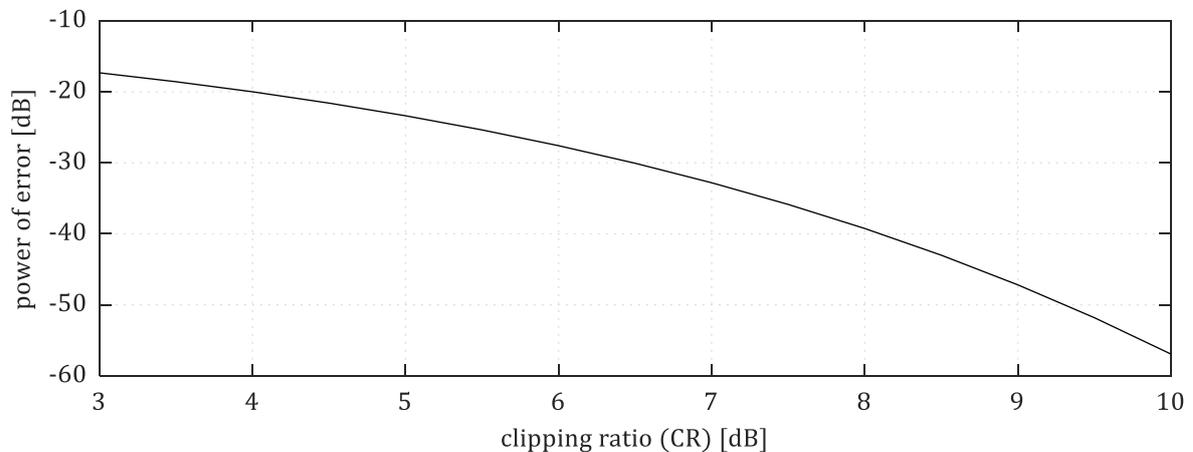


Figure 4.10: The power of error resulting from peak limitation reduces with increasing clipping ratio (CR).

When looking at the relation of the EVM prediction based on the power of the error, from Figure 4.10, to the CR in a linear scale, presented in Table 4.2, about 5 dB slope of the error power is clearly visible.

Table 4.2: The relation between the linearly scaled clipping ratio (CR) and the EVM prediction.

$10^{CR/10}$	1	2	3	4	5	6	7	8	9	10
EVM [dB]	-12.5	-17.4	-22.5	-27.7	-32.8	-37.7	-42.6	-47.5	-52.2	-56.9
EVM [%]	23.7	13.6	7.5	4.1	2.3	1.3	0.7	0.4	0.2	0.1

The per antenna EVM is estimated based on a linear regression model with a single explanatory variable, representing the estimated power of error from the relation visualized in Figure 4.10 for a given CR and a constant:

$$\text{EVM}_{\text{ant}}(\text{CR}) = \delta_0 + \delta_1 f(\text{CR}) \quad (4.17)$$

where the value of the constant term δ_0 is expected to depend on the used peak limitation approach, with the EVM estimate scaled:

- Down for HC, since some of the IMPs fall at frequencies not occupied by SCs of the 5G NR carrier, hence not contributing to the modulation error.
- Up for ICF, since the IMPs that fall outside of the 5G NR carrier after the clipping block are filtered out, leading to additional IMPs (generated in the following iterations of clipping operation) falling at the used SCs.

In the end, what matters is the way the per antenna error radiates and is received by the UEs. The per user error prediction models, introduced in Section 4.4.2, reflect that crucial aspect of directional radiation patterns of nonlinear distortion.

4.4.2. Per User Error

The expectation for the nonlinear distortion, based on the behavior described in Section 4.3.3, is that it partially radiates to the same directions as the wanted signal. The power of that user-directed portion of the IMPs is generally expected to reduce when the number of UEs (or more generally, the number of directions with significant power) is increased. This is visible in Figure 4.11, where the ratios of powers of the user-directed and other spatially filtered third-order IMPs to the sum power of all the third-order IMPs are reflected, based on U from (4.16).

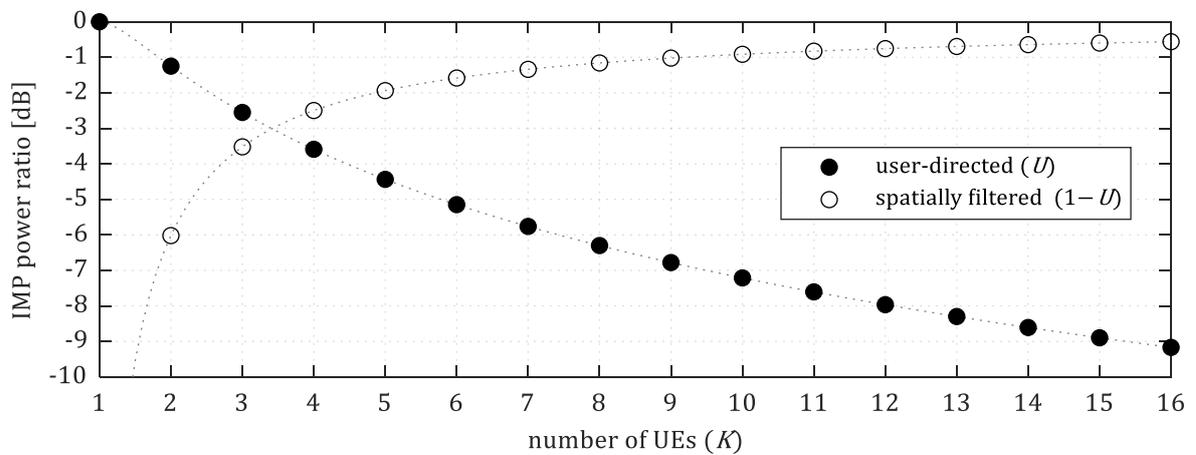


Figure 4.11: The theoretical power ratios of the user-directed and other spatially filtered third-order intermodulation products (IMPs) to all the third-order IMPs.

Note that higher-order IMPs, which are not covered by the U ratio, can also radiate to the UE directions. While their impact is less relevant, since their powers are expected to be significantly lower than that of the third-order IMPs, the expectation is that the power level of the user-directed IMPs is in practice somewhat lower than presented in Figure 4.11.

When relying on (4.16), the relation between the amount of the user-directed IMPs and the other spatially filtered IMPs does not depend on the number of antennas M . Hence, the error caused by nonlinear distortion is not expected to vanish when $M \rightarrow \infty$. In fact, for a transmission to 1 UE in a free-space LoS channel the per user error is going to be on a par with the per antenna error, irrespective of the number of antennas used for transmission [21], since no spatially filtered IMPs are expected to be generated.

An EVM prediction based on (4.16), assuming equal power allocation to layers and SCs, could be defined as

$$\text{EVM}_{\text{UE}} = f(\text{EVM}_{\text{ant}}, K) = \text{EVM}_{\text{ant}} \cdot \sqrt{U} = \text{EVM}_{\text{ant}} \cdot \frac{\sqrt{(2K-1)}}{K} \quad (4.18)$$

where EVM_{ant} represents the per antenna EVM, in a linear scale, assuming equal CR (resulting from using PAs with the same peak power capabilities in all the antennas) and roughly equal power of the precoding weights over antennas (which could be achieved by a similar approach to FA-based beamforming proposed in [48]) for simplicity. Note that the estimate of the per user EVM_{UE} is equal for all the K UEs in this case.

The EVM performance calculated based on (4.18) is expected to lead to overoptimistic estimates in practice. In fact, there are other factors that influence the EVM performance, beyond the number of UEs. These additional aspects are explored in following subsections, starting with the number of antennas, going through the unfavorable sets of angular directions (covering also a definition of the category of worst-case sets), ending with the limited angular resolution of radiation from the array (for both linear and rectangular arrays).

4.4.2.1 The Best-Case Prediction Model: Reflecting the Dependence on the Number of Antennas

When a prediction model based solely on the number of UEs is considered, as in (4.18), the expectation is that the EVM improves with increasing number of UEs K , seemingly without bound. In fact, for a relatively large K , the EVM seems to improve by up to about 3 dB with each doubling of K , since (4.16) starts to be relatively well approximated by $2/K$ when K is increasing. However, some dependence on the number of antennas M , related with the way the power is allocated to UEs, could be expected.

When the UEs are spatially multiplexed, the sum output power is shared between them. While the array gain increases with M , the per UE output power decreases with K (when operating in the high-power mode [25]). This suggests that, at some point, even uniformly radiating distortion should limit the EVM improvement. Such limitation is visualized in Figure 4.12, assuming an equal power allocation between spatially multiplexed UEs. When the number of UEs exceeds a limit for a given number of antennas, the EVM can be expected to start increasing since the per UE power is getting closer to the power level of uniformly radiated distortion.

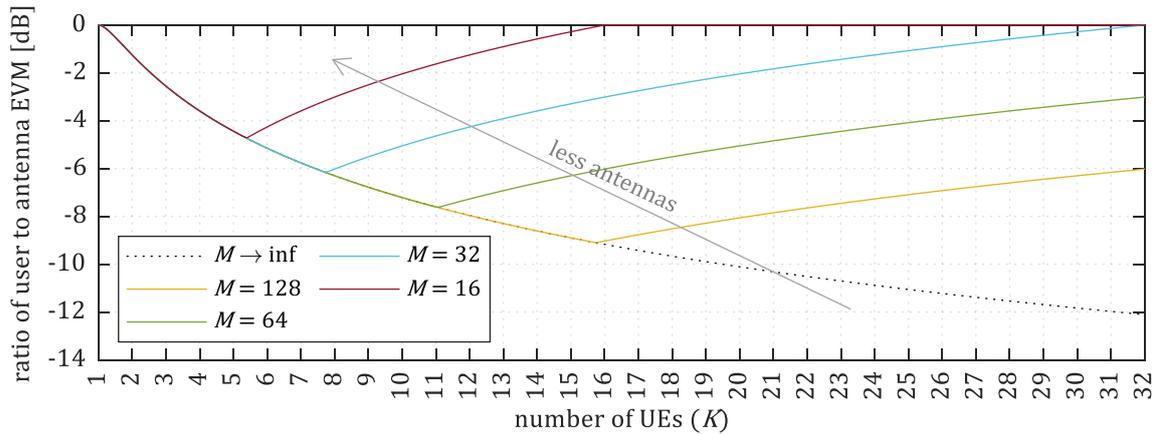


Figure 4.12: The uniformly radiated uncorrelated noise limits the reduction of EVM when increasing the number of UEs K , depending on the number of antennas M .

Therefore, a model of EVM performance should reflect the potential performance floor related with uniformly radiating distortion, reducing the per UE SDR when the signal power is split between UEs. The EVM calculation in (4.18) can be extended, to cover the dependence on the number of antennas M , as follows:

$$\begin{aligned} \text{EVM}_{\text{UE}} = f(\text{EVM}_{\text{ant}}, K, M) &= \text{EVM}_{\text{ant}} \cdot \sqrt{\max\left(U, \frac{K}{M}\right)} \\ &= \text{EVM}_{\text{ant}} \cdot \sqrt{\max\left(\frac{(2K-1)K}{K^2}, \frac{K}{M}\right)}. \end{aligned} \quad (4.19)$$

It is not excluded that in the best-case for nonlinear distortion the EVM performance improvement depending on the number of antennas may be larger than in Figure 4.11, since the angular directions of UEs may be selected in a way that the IMPs radiate mostly away from the UEs. However, it is expected that typically (when many sets of angular directions for the wanted signal are considered) the limitation of EVM, based on the assumption of uniformly radiated distortion should hold.

The influence of the actual angular directions of UEs on the EVM performance is considered next, with a focus on unfavorable combinations of steering angles.

4.4.2.2 Unfavorable Combinations of Steering Angles

While the ratio between the user-directed IMPs and the spatially filtered IMPs depends on the number of UEs, some of the spatially filtered IMPs (which radiation directions are resulting from precoding weights) may actually radiate in the same directions as the wanted signal. A more detailed explanation with examples is provided in the Appendix Section B.2.

For a third-order nonlinearity, two classes of spatially filtered IMPs are generally expected to radiate in distinct directions:

- $2A-B$: where A and B represent components of IMPs formed by the precoded data from different UEs;
- $A+B-C$: where C represents a component from different UE than A and B .

Note that in the $A+B-C$ class, when A and B are components associated with the same UE (i.e., are generated by different SCs of the same MIMO layer), the resulting radiation direction is the same as for $2A-B$. This results from the same phase gradient used for all the SCs in a given MIMO layer, to precode data for transmission in free-space LoS channels.

The directions of radiation of IM beams depend on the array geometry, its orientation in space, and definition of the steering angles. While a detailed description of these aspects is provided in Appendix A and B, examples of the relation of phase gradients of IMPs and the resulting steering angles are provided in the following subsection for azimuth steering in a horizontally-oriented ULA with half-wavelength spacing.

Uniform Linear Arrays with Horizontal Beam Steering

Assuming ULA with no elevation steering, the phase gradient for the k -th UE defined as $\Delta\gamma_k$, there are certain combinations of $\Delta\gamma \in \Gamma$ generated by the nonlinearity (i.e., $2A-B$ and $A+B-C$ types for a third-order nonlinearity), where Γ is the set of K phase gradients used by the per UE precoding weights, that may result in a subset of IMPs radiating to a subset of the azimuth steering directions $\phi_k \in \Phi$ associated with UEs. This situation happens when either $(2\Delta\gamma_A - \Delta\gamma_B + n2\pi) \in \Gamma$ or $(\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C + n2\pi) \in \Gamma$ for any integer n , for $2A-B$ and $A+B-C$ type of IMPs, respectively. A more detailed description is provided in Appendix B.

For example, considering 3 UEs a situation that IMP generated by the interaction of the 1st and 2nd UE lead to IMP falling at the 3rd UE location happens when either

$$\Delta\gamma_C = (2\Delta\gamma_A - \Delta\gamma_B + n2\pi) \quad (4.20)$$

$$\Delta\gamma_C = \frac{(\Delta\gamma_A + \Delta\gamma_B + n2\pi)}{2}. \quad (4.21)$$

When analyzing for $n = 0$, (4.20) and (4.21) simplify to

$$\Delta\gamma_C = (2\Delta\gamma_A - \Delta\gamma_B) \quad (4.22)$$

$$\Delta\gamma_C = \frac{(\Delta\gamma_A + \Delta\gamma_B)}{2}. \quad (4.23)$$

In both above cases it is apparent that the phase gradient for C can be selected such that there are equal distances between the phase gradients for all 3 components. This is possible with either a wider distance and $\Delta\gamma_B < \Delta\gamma_A < \Delta\gamma_C$ for (4.22) or a narrower distance and $\Delta\gamma_A < \Delta\gamma_C < \Delta\gamma_B$ for (4.23).

To demonstrate that a subset of spatially filtered IMPs is expected to radiate to the UE directions when at least 3 phase gradients are at equal distances from each other, a comparison of theoretical radiation patterns (of both the wanted signal and third-order IMPs) for ULA with a half-wavelength spacing between antennas, is visualized in Figure 4.13 for two sets of 3 phase gradients that differ only in selection of one of the phase gradients.

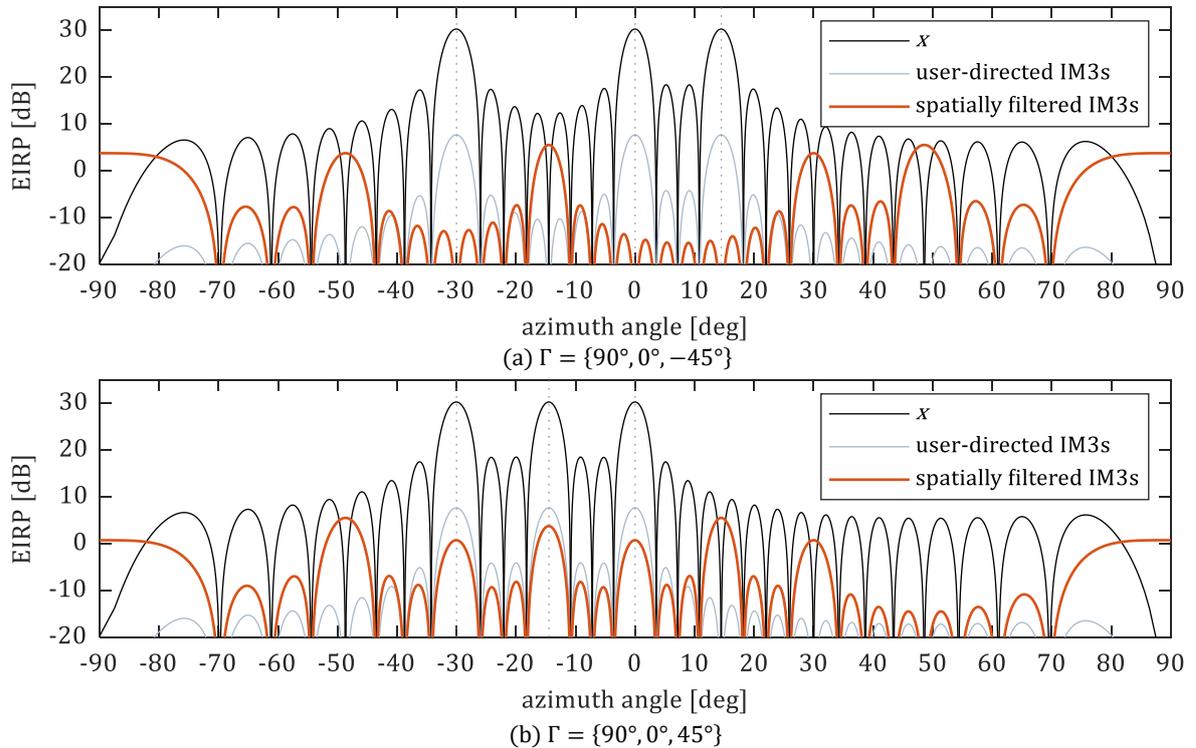


Figure 4.13: Expected radiation patterns of the wanted signal x as well as the spatially filtered and the explicitly user-directed third-order intermodulation products (IM3s), for two different sets Γ of phase gradients $\Delta\gamma$.

In case the set of phase gradients does not follow the constraint of equal distances the spatially filtered IMPs radiate to distinct directions that are different than for the wanted

signal, as clearly visible in Figure 4.13(a). While, when the third phase gradient is modified in a way that equal distance of 45° is achieved, some of the spatially filtered IMPs form beams in the UE directions, as can be observed in Figure 4.13(b).

Based on the knowledge of angular directions of the LoS paths for each of the UEs, the distances between the phase gradients could be verified in order to identify what portion of the spatially filtered IMPs could be expected to be radiated towards the UEs. This could be used as an additional computation step, modifying the EVM prediction from (4.19) by taking into account additionally the sets of azimuth steering directions Φ , elevation steering directions Θ , and both horizontal and vertical spacing between antennas in the array.

An extension of the analysis of unfavorable sets of angles, covered in Section 4.4.2.3, reveals the existence of worst-case combinations of steering angles, resulting in all the spatially filtered IMPs radiating towards UEs.

4.4.2.3 Worst-Case Combinations of Steering Angles

Combinations of steering angles leading to radiation of all the spatially filtered IMPs in the same directions as the wanted signal were analytically identified, by extending the analysis of the unfavorable sets of directions. These worst-case combinations are expected to occur irrespective of the number of spatially multiplexed UEs. This happens for a specific subset of the category of equal distances between phase gradients (which was demonstrated in Section 4.4.2.2 to lead to radiating subset of spatially filtered IMPs to the angular directions existing in the wanted signal), as explained in following subsections based on ULA and URA examples.

Uniform Linear Arrays with Horizontal Beam Steering

With $K = 2$, radiation of all IMPs towards UEs happens when either of the following is true

$$2\Delta\gamma_A - \Delta\gamma_B + n2\pi = \Delta\gamma_A \quad (4.24)$$

$$2\Delta\gamma_A - \Delta\gamma_B + n2\pi = \Delta\gamma_B \quad (4.25)$$

where n is an integer, which leads to:

$$\Delta\gamma_B = \Delta\gamma_A + n2\pi \quad (4.26)$$

$$\Delta\gamma_B = \Delta\gamma_A + n\pi. \quad (4.27)$$

Clearly, with $n = 0$, either (4.26) or (4.27) is true only when $\Delta\gamma_B = \Delta\gamma_A$. This would be the case either if the same phase gradient was applied to both UEs, which is not possible without FDMed UEs, or if both A and B components are from the same MIMO layer, which maps to the definition of the user-directed IMPs.

For $n = \pm 1$, the same logic applies to (4.26), however (4.27) is fulfilled by

$$\Delta\gamma_B = \Delta\gamma_A \pm \pi. \quad (4.28)$$

Therefore, for a half-wavelength spaced ULA, for any azimuth steering angle ϕ_1 there exists a single (provided that $\phi_k \neq \pm\pi$) different steering angle ϕ_2 , related with the first one by (4.28), which results in radiating all the spatially filtered third-order IMPs in the directions of both UEs, when $K = 2$.

With $K = 3$, more conditions must be met than for 2 UEs for all the spatially filtered IMPs to radiate only to the same angular directions as the wanted signal. Namely no 2A–B and A+B–C combination should be able to generate a new azimuth angle. This means that the following must be fulfilled:

$$2\Delta\gamma_A - \Delta\gamma_B + n2\pi = \{\Delta\gamma_A, \Delta\gamma_B, \Delta\gamma_C\} \quad (4.29)$$

$$\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C + n2\pi = \{\Delta\gamma_A, \Delta\gamma_B, \Delta\gamma_C\} \quad (4.30)$$

for all $n = (0, \pm 1, \pm 2, \dots)$, and all combinations of A , B , and C .

The fulfillment of (4.30) requires the phase gradient to be equidistantly spaced from each other, as demonstrated in the previous subsection.

Upon a closer examination of (4.29), it can be observed that in order for a combination of $n_1 = 0$ and $n_2 = \pm 1$ for $\Delta\gamma_A$ and $\Delta\gamma_B$ to give the same $\Delta\gamma_C$ the following must be fulfilled:

$$2\Delta\gamma_A - \Delta\gamma_B = 2\Delta\gamma_B - \Delta\gamma_A \pm 2\pi. \quad (4.31)$$

Solving (4.31) for $\Delta\gamma_A$ leads to:

$$\Delta\gamma_A = \Delta\gamma_B \pm \frac{2}{3}\pi. \quad (4.32)$$

Similarly, looking for a combination of $n_1 = \pm 1$ and $n_2 = \pm 1$ for $n_1 \neq n_2$ in (4.29) the following must be fulfilled:

$$2\Delta\gamma_A - \Delta\gamma_B - 2\pi = 2\Delta\gamma_B - \Delta\gamma_A + 2\pi \quad (4.33)$$

$$2\Delta\gamma_A - \Delta\gamma_B + 2\pi = 2\Delta\gamma_B - \Delta\gamma_A - 2\pi. \quad (4.34)$$

Solving both (4.33) and (4.34) for $\Delta\gamma_A$ and combining these solutions gives:

$$\Delta\gamma_A = \Delta\gamma_B \pm \frac{4}{3}\pi. \quad (4.35)$$

In fact, both (4.32) and (4.35) lead to the same set of steering angles (defined by a set of phase gradients $\Gamma = \{\Delta\gamma_A, \Delta\gamma_A + 2\pi/3, \Delta\gamma_A + 4\pi/3\}$), due to a 2π periodicity of the steering angle calculation in (4.15).

The pattern visible for 2 and 3 UEs with equal differences of the phase gradients of $2\pi/K$, leading to radiating all the spatially filtered third-order IMPs to the azimuth angles existing in the wanted signal, is expected to hold also for $K > 3$.

Examples of worst-case combinations of steering angles in URAs are presented next.

Uniform Rectangular Arrays

For a transmission from URA, worst-case sets of phase gradients resulting in radiating all the IMPs in the direction of UEs can also be expected, similarly to ULA. Three classes of scenarios, demonstrated with examples in Figure 4.14, were identified here to lead to the worst-case performance, with all spatially filtered IMPs radiating towards UEs.

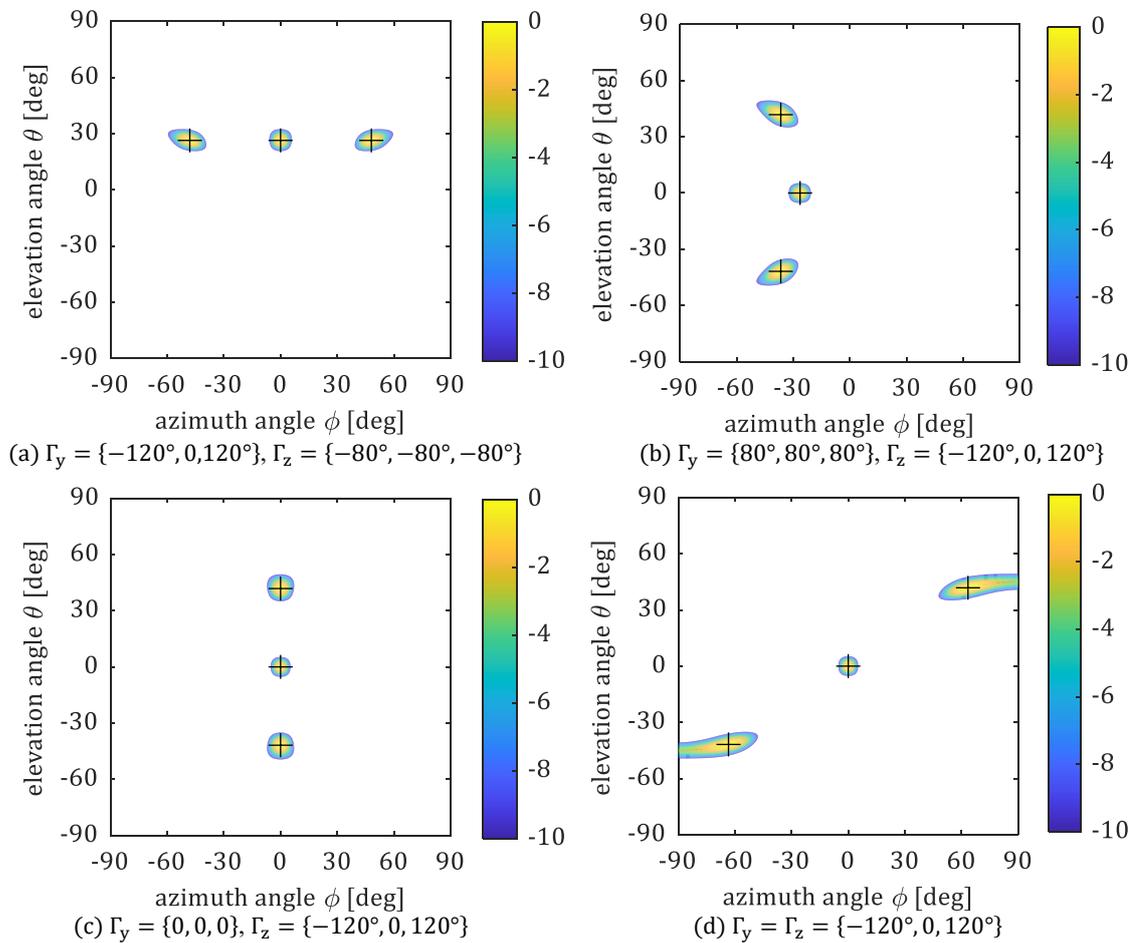


Figure 4.14: Normalized radiation patterns, excluding values below -10 dB, for the expected spatially filtered third-order intermodulation products of 16×16 URA for 4 different sets of phase gradients (with beams pointing directions for the wanted data signal visualized with crosses).

The examples of the worst-case sets of user locations, presented in Figure 4.14 can be categorized as follows:

1. Horizontal phase gradients spaced by $2\pi/K$ and vertical phase gradients equal (leading to equal elevation steering angles). This is exemplified in Figure 4.14(a).
2. Vertical phase gradients spaced by $2\pi/K$ and horizontal phase gradients equal to each other. This generally leads to nonequal azimuth steering angles, as exemplified in Figure 4.14(b), except when the horizontal phase gradient is 0° equal azimuth steering angles of 0° , as visualized in Figure 4.14(c).
3. Both horizontal and vertical phase gradients spaced by $2\pi/K$, as exemplified in Figure 4.14(d).

The analysis of the radiation directions of the wanted signal, resulting in distinct directions of IMP radiation, suggests that consideration of these spatial characteristics of the transmission scenario may improve the EVM prediction performance of the best-case model defined in Section 4.4.2.1. Alternatively, the precoding matrix could be used directly, to calculate the implicit precoding vectors applicable to the spatially filtered IMPs. This is a basis for the precoding-based prediction model, described in Section 4.4.2.4, which additionally allows for consideration of the impact of the limited angular resolution of radiation from the antenna array.

4.4.2.4 The Precoding-Based Prediction Model: Addressing the Limited Angular Resolution of the Array

The spatial resolution of the array is not infinite. Hence, even if the main lobes of IMPs do not point directly to the same angles as the wanted signal, part of their radiation energy can actually be received by UEs. This effect is not visible when relying solely on the information about the angles at which the main lobes of IMPs are pointing, as in Sections 4.4.2.1–4.4.2.3.

Following boundary cases of spatial scenarios were identified for MU-MIMO transmission in LoS channels, when neglecting the effects of a finite angular resolution:

- Best-cases, in which no spatially filtered IMPs are radiated in the directions of UEs, as assumed in the simplified best-case prediction model in Section 4.4.2.1.
- Worst-cases, in which the radiation pattern of distortions follows the wanted signal (as widely expected for the single UE case [21]), hence there are no IMPs radiating in directions other than the wanted signal, as demonstrated in Section 4.4.2.3.

Between these boundary cases there is a spectrum of scenarios of spatial distribution of UEs, for some of which a portion of the spatially filtered IMPs will radiate to the directions of users. This was shown for unfavorable combinations of angles, in Section 4.4.2.2.

In practice however, the finite angular resolution of the radiation from the antenna array results in a portion of the energy of the spatially filtered IMPs to be received by the UEs, even when there are no spatially filtered IMPs that radiate directly to the UE locations. This is especially relevant when considering how the antennas are positioned in the array, since the granularity in azimuth and elevation is separately determined by the array aperture. With a given number of antennas M , a linear array will provide a finer angular resolution than a rectangular array in one of the steering dimensions, at the cost of no steering capability in the other dimension.

The impact of the limited angular resolution can be demonstrated with an example. A comparison of the normalized radiation patterns (excluding values below -10 dB, for clarity) of the spatially filtered third-order IMPs is visualized in Figure 4.15, for square 32×32 and 8×8 URAs for transmission to two sets of 4 UEs with randomly selected angular directions.

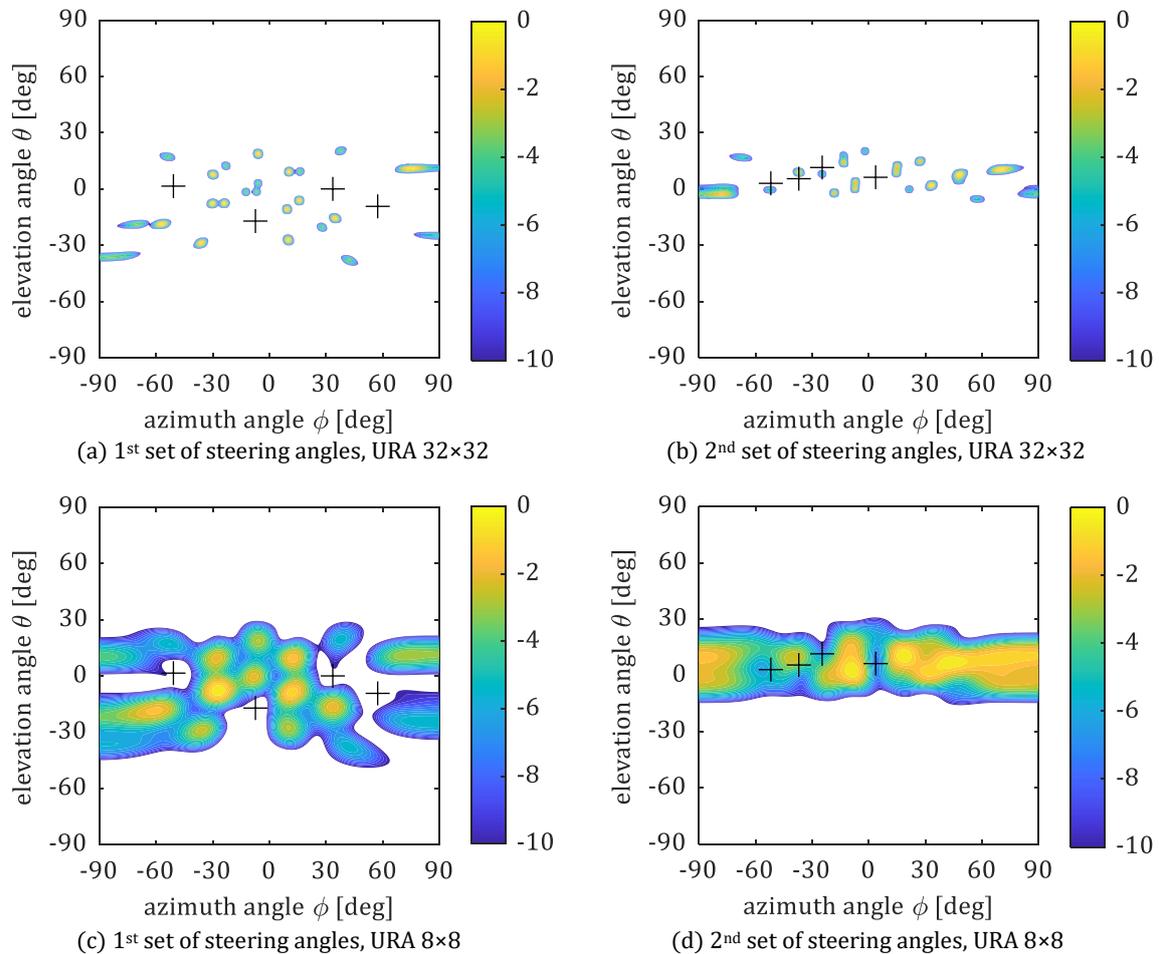


Figure 4.15: Comparison of normalized radiation patterns, excluding values below -10 dB, of spatially filtered third-order intermodulation products radiated from 32×32 and 8×8 rectangular arrays, for 2 sets of steering angles for 4 UEs (with directions of UEs marked with crosses).

In the examples visualized in Figure 4.15 the UEs locations, marked with crosses, were randomly selected from azimuth range constrained to $\pm 60^\circ$ and elevation range limited to $\pm 20^\circ$. Clearly, both the geometry of the array as well as the precoding (resulting from the UE locations) matters, since only in case of precoding based on the 2nd set of angles for 8×8 URA in Figure 4.15(d) a direct overlap of the distortion and the UE locations can be observed. Note that it does not mean that in the other cases, visualized in Figure 4.15(a)–(c), there was no distortion power radiated towards UEs. It means that the distortion power radiated towards UEs was more than 10 dB lower than the power in the highest main lobe of the distortion radiation pattern, leading to potentially receiving only a minor portion of the spatially filtered IMPs.

In the end, the per user error depends on what portions of the spatially filtered IMPs are received by the respective UEs. This, on the one hand, depends on the structure of the array (i.e., number of antennas, array geometry, and array element radiation pattern) used by the base station for transmission. Since that structure, together with the precoding weights applied to the wanted signal, define the radiation pattern of the transmitted distortion. On the other hand, the characteristic of the channel between the base station and the UE determines what portion of that radiated distortion is actually received.

Therefore, a precoding-based EVM performance prediction model is proposed, to capture the aspect related with limited angular granularity and to cover the general cases of precoding in free-space LoS scenarios (i.e., not only the specific cases of beam steering).

First, the per UE precoding vectors $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ applied to the data samples s_k , according to (4.3), are used to calculate the implicit precoding vectors $\mathbf{u}_i \in \mathbb{C}^{M \times 1}$ effectively applicable to the i -th third-order IMP, in line with the approach in [2]:

$$\mathbf{u}_i = \mathbf{w}_A \odot \mathbf{w}_B \odot \mathbf{w}_C^* \quad (4.36)$$

where \odot denotes the element-wise product of vectors, while the indices A , B , and C select the transmission layers from $\{1, \dots, K\}$, where K is the number of UEs. Note that some of these precoding vectors \mathbf{u}_i represent the category of explicitly user-directed IMPs, with $\mathbf{u}_i = \mathbf{w}_k$ due to specific selection of A , B , and C (e.g., when $A = B = C$), same as in the case of beam steering (explained in more detail in the Appendix Section B.2) based on the relations between the per UE phase gradients.

The explicitly user-directed IMPs are considered here separately, in line with the simplified EVM calculation in (4.18). Hence, only the subset of precoding vectors \mathbf{u}_i that

represent spatially filtered IMPs is explicitly covered in the next step, where the expectation of the portion of the power of spatially filtered IMP received by UE is evaluated.

Note that in case of unfavorable combinations of angles some of \mathbf{u}_i may be equal to some of \mathbf{w}_k , leading to radiation of a subset of the spatially filtered IMPs to the UE locations, as explained in Section 4.4.2.2. This in the worst-case may result in receiving full power of the spatially filtered IMPs, as revealed in Section 4.4.2.3.

Assuming the channel responses are normalized as $\mathbb{E}\{\|\mathbf{h}_k^T\|^2\} = 1$ and precoding vectors are normalized as $\mathbb{E}\{\|\mathbf{w}_k\|^2\} = M/K$, the received portion of the spatially filtered IMPs can be described as

$$\rho_k = \frac{\sum_{i=1}^N |\mathbf{h}_k^T \mathbf{u}_i|^2}{(M/K)(1-U)} = \frac{K}{M(1-U)} \sum_{i=1}^N |\mathbf{h}_k^T \mathbf{u}_i|^2 \quad (4.37)$$

where $\mathbf{h}_k^T \mathbf{u}_i$ is the gain of the effective precoded channel for the i -th IMP, N is the number of spatially filtered IMPs, M/K is the expected gain of the precoding vector per UE in case of equal power allocation, $(1-U)$ is the spatially filtered portion of all the third-order IMPs.

Note that $\rho_k \in [0,1]$ in case of equal power allocation to UEs. With $\rho_k = 0$ representing a situation that no power of the spatially filtered IMPs is received by the k -th UE, while $\rho_k = 1$ reflecting the worst-case of receiving full power of the spatially filtered IMPs, proportionally by each of the K UEs.

A precoding-based EVM performance prediction model is proposed to be based on the simplified per user EVM calculation from (4.18), extended with the dependence on both the precoding matrix \mathbf{W} and the channel matrix \mathbf{H} , reflected by the portion of the spatially filtered IMPs received by the k -th UE defined by (4.37). Instead of having a single EVM_{UE} estimate applicable to all UEs like in the best-case EVM prediction model from (4.19), this results in calculation of per UE EVM_k separately for each of the K UEs, as follows:

$$\begin{aligned} \text{EVM}_k &= f(\text{EVM}_{\text{ant}}, K, M, \mathbf{W}, \mathbf{H}) = \text{EVM}_{\text{ant}} \sqrt{U + \rho_k(1-U)} \\ &= \text{EVM}_{\text{ant}} \sqrt{U + \frac{K}{M} \sum_{i=1}^N |\mathbf{h}_k^T \mathbf{u}_i|^2}. \end{aligned} \quad (4.38)$$

Note that in the worst-case (i.e., when $\rho_k = 1$) the per user EVM_k is equal to the per antenna EVM_{ant} , irrespective of the ratio U (of the explicitly user-directed IMPs to all the IMPs).

Note that the explicit constraint related with the limiting aspects of uniformly radiating distortion, represented in the best-case prediction model in (4.19) by the dependence on the

number of antennas, is not required in the precoding-based prediction model in (4.38), since this aspect is implicitly covered by calculation of the portion of the received spatially filtered IMPs ρ_k in (4.37).

4.5. Expected Error Vector Magnitude Performance for Quantization

The bitwidth and the sample rate determine the absolute level of the quantization noise power. However, the ratios between the powers of the per UE wanted signals and the quantization noise power depend on the precoding weights. The power allocation strategy, selected out of the options introduced in Section 4.1.1, is an important factor modifying the impact of the quantization noise over UEs and SCs.

The expectation is that the quantization noise, modeled with Gaussian noise uncorrelated between antennas, will radiate with the antenna array element pattern, as demonstrated in Figure 4.16. Hence, when the array gain increases with increasing the number of antennas, the per user EVM resulting from the quantization noise is expected to decrease. This is confirmed by about 18.1 dB difference visible in the power radiated at the 30° steering angle for x and d in Figure 4.16, resulting from the array gain for 64 antennas. This is a very different behavior than expected for a peak power limitation, as covered in Section 4.4.

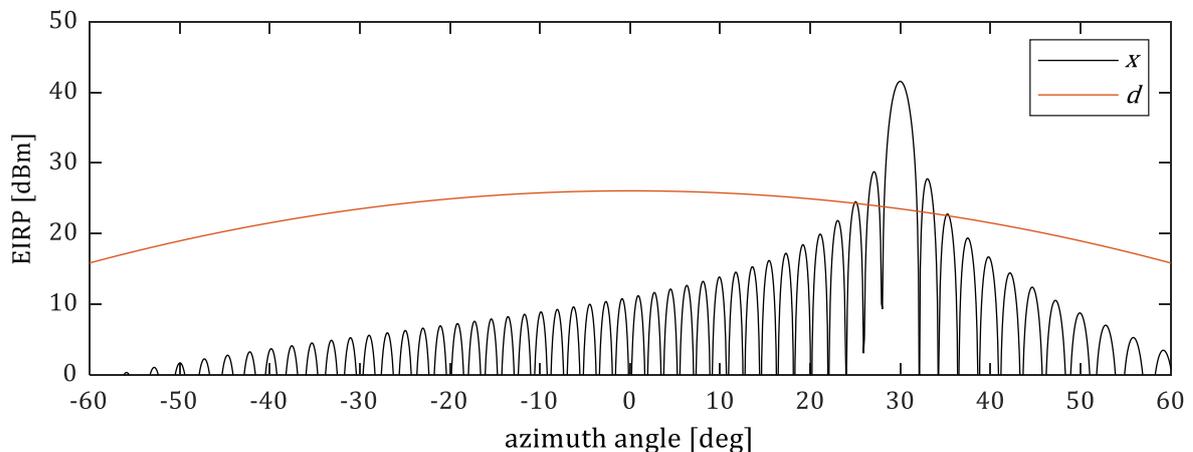


Figure 4.16: Radiation patterns of the wanted signal x and distortion d for uniform linear array (ULA) with 64 antennas, using directional antenna elements with 8 dBi, for steering to 30° .

When it is assumed that the power of distortion d scales proportionally with the power of the wanted signal x per antenna (as in [1]), it is trivial to estimate the relation between increasing the number of antennas M and the resulting reduction in EVM at the user for the case of 1 UE in a free-space LoS channel. That reduction is approximately equal to the difference between an array gain of $10\log_{10}M$ decibel applicable to the coherently

combining wanted signal x and an array gain of 0 dB applicable to the incoherently combining distortion d .

In the MU-MIMO transmission case, the array gains applicable to both x and d are the same as in the MISO transmission to 1 UE. However, the power of the part of the wanted signal that is precoded to radiate to the k -th UE is only a portion of the power of x (which is constructed from data symbols precoded for K UEs). Hence in MU-MIMO, the power allocated per UE is $p_k < P$ and is equal to $p_k = P/K$ in case of an equal distribution of the sum output power P . Therefore, compared to the MISO transmission case, the ratio between the power of the portion of the wanted signal x transmitted to the k -th UE and the power of the uncorrelated distortion d will reduce (by K in case of an equal power distribution), leading to increasing per UE EVM with increasing number of UEs. This applies in both the high-power and the low-power operating mode of the Massive MIMO base station, as explained already in [62].

The validity of the expectation for the EVM performance with quantization noise is explored in Appendix Section D.4 for equal power distribution over UEs and in Section 4.7.5 for unequal power distribution.

4.6. Methods

A Matlab-based link-level simulation (LLS) environment, described in more detail in Appendix C, was developed and used to analyze the impacts of the residual TX impairments on the EVM. That simulator reflects the system model introduced in Section 4.1 and the models of TX distortion described in Section 4.2. It was already demonstrated in [62], that the developed simulator provides a comparable relation between the block error rate (BLER) and the SNR to the performance visible in similar simulations in 3GPP studies [68]. This section dives into details of how that LLS was used to attain the simulation results presented later, in Section 4.7.

The simulation results obtained in LLS are compared with numerical simulations based on the analytical model, proposed in Section 4.4.2.4, using precoding weights generated in Monte Carlo trials with varying UE locations in free-space LoS channels.

4.6.1. Antennas

Single uni-polarized omni-directional antennas are used for reception at the user side, for each of the K UEs.

While at the base station side, uniform array with M uni-polarized antennas, spaced at half the wavelength λ at 3.5 GHz are used for transmission (targeting the n78 NR operating band [65], covering 3.3–3.8 GHz). The half-wavelength distance between elements in the array is as typically selected to avoid spatial aliasing and limit the mutual coupling between the antennas in the array.

The simulation results are obtained primarily with horizontally oriented ULA. The impacts of elevation steering are demonstrated with URA, which is also used to cover the limited spatial resolution available in Massive MIMO base stations (especially when relatively small number of columns is assumed, in alignment with the array structures used in the initial deployments [11]).

Omni-directional antennas are used in case of ULA, while URA is formed by an array of isotropic sources.

4.6.2. Resource Allocation

A single 5G NR carrier (using a CP-OFDM waveform, which is the only option for DL in 5G NR) with 20 MHz channel bandwidth and 15 kHz spacing between SCs is utilized. This gives 1272 SCs that can be used for DL transmission (including pilot signals) in a given slot. All the REs are allocated to all the UEs, without co-scheduling UEs in the frequency domain.

4.6.3. Channels and Precoding

Spatially correlated clustered delay line (CDL) channel models, defined by 3GPP in [69, Sec. 7.7.1] are used to model the MISO channels (either multi-path LoS or NLoS) between the base station and each of the UEs. While these models can be used to represent a variety of environments, with the selected frequency band and an RMS delay spread of 100 ns, the focus is primarily on UMa and UMi deployments. The coherence block, with a time-invariant and frequency-flat channel response, is assumed to cover 1 SC in frequency and 1 slot in time. Hence, the fading realizations are changing between slots, but not during slots. Additionally, deterministic free-space LoS channels are used for reference.

A perfect channel knowledge at the base station side is assumed. For each slot, a per SC precoding matrix \mathbf{W} is calculated based on the channel matrix \mathbf{H} . The ZF precoding is used to make the impact of the intra-cell interference negligible, which allows to focus on the effects of TX distortion. The calculated precoding weights are scaled based on the power allocation schemes introduced in Section 4.1.1. By default, an equal power allocation between UEs is used, similarly like in [1].

4.6.4. Equalization and Calculation of the Error Vector Magnitude

At the UE side, a practical channel estimation based on the pilot signals (i.e., the DMRS) is used. The signal received by the UE is passed through a minimum mean-squared error (MMSE) equalizer. This is a deviation compared to 3GPP conformance verification, where a simpler equalizer is used for the EVM measurement [66, Sec. L.6].

The EVM calculation is done as RMS over all REs in a slot, which means that per slot there is a single EVM per antenna or a single EVM per user.

4.6.5. Repeatability of Simulation Results

The frequency-domain data samples s_k are generated based on pseudorandom bitstreams per UE, passed through a Downlink Shared Channel (DL-SCH) encoder and a QAM modulator. When testing multiple slots, the input data changes between slots. However, when comparing behavior in different channel models, different spatial scenarios, or different peak power limits, the same set of input data is used. This removes the degree of freedom related with the randomness of the input data.

4.6.6. Rejection of Selected Simulation Results

Spatial multiplexing of UEs that are angularly too close to each other is avoided, in order to isolate the impact of distortion. Specifically, degradation of the performance of the ZF precoding (resulting in reduction of the power of the wanted data signal radiated to UEs) is avoided. This is achieved in two steps. The first step is a restriction on the minimum angular difference of UE location (e.g., about 3° for 32 antennas) for the random selection of UEs in the Monte Carlo trials. The second step, taken after obtaining the simulation results, is rejection of results in which the received power of the wanted signal is more than 1 dB lower than expected with MR precoding.

4.7. Simulation Results and Discussion

The impact of the residual TX impairments on the EVM performance was explored using both LLS and numerical simulations (using precoding matrix to generate expectations of EVM, according to the precoding-based prediction model defined in Section 4.4.2.4).

The feasibility of utilizing the per antenna EVM prediction (based on the power of the clipping error obtained for circularly symmetric complex Gaussian signal, as proposed in Section 4.4.1) to describe the EVM performance of ICF and HC is established first. The relation of the per antenna EVM and the peak power limit obtained during this analysis, as

captured in Section 4.7.1, is used as a reference for the per user EVM considered in the following verification steps.

The key for confirming the credibility of the per user error models defined in Section 4.4.2, is validation of the EVM performance in the worst-case sets of angular directions of UEs (identified in Section 4.4.2.3). This is done in Section 4.7.2, for both deterministic and frequency-selective LoS channels.

The EVM performance obtained for the best- and worst-case sets of angular directions (defined in Section 4.7.2) provides boundary conditions for the range of EVM expectations in the general case of randomly placed UEs, which is explored in Section 4.7.3 for both ULA and URA. This analysis considers the impact of the number of spatially multiplexed UEs, number of antennas, and antenna geometry. The primary purpose is to establish the prediction accuracy offered by the models proposed in Section 4.4, to evaluate their efficacy for providing valid expectations for the distortion power received by UEs.

The EVM performance in frequency-selective NLoS scenarios, covered in 4.7.4, complements the expectations for LoS channels. This enhances the validity of the way the impacts of the residual TX impairments on the achievable DL data rates are explored later, in Chapter 5.

Finally, the impact of the power allocation strategy is explored in Section 4.7.5, for both quantization noise and peak power limitation, by considering the relation between SNR and EVM, forming a bridge to the analysis of the impact of distortion on the data rate performance, covered in Chapter 5.

4.7.1. Accuracy of the per Antenna Error Model for Peak Power Limitation

Both of the proposed model of per user error rely on the estimate of the per antenna EVM, based on the model described in Section 4.4.1. Hence, the validity of the underlying assumption that an EVM prediction based on a clipped complex Gaussian signal can be representative of the EVM resulting from the peak limitation of a 5G NR signal is tackled.

To judge the accuracy of the per antenna EVM model, simulations for both the ICF and the HC were performed, covering a range of CRs from 3 to 10 dB. For each of these considered peak power limitation approaches, the same set of randomly generated input data covering 100 slots was used for each of the CR levels.

The comparison of the EVM estimated from the prediction model with the EVM statistics calculated based on the simulation results is presented in Figure 4.17. The solid lines represent the RMS of per slot EVM values, calculated over all the simulated slots, while

the filled ranges span a distance of 1 standard deviation from these RMS values. For each tested CR level, the EVM for the ICF was higher and for the HC (representing here the saturation in a linearized PA) was lower than prediction, which is in line with the expectations described in Section 4.4.

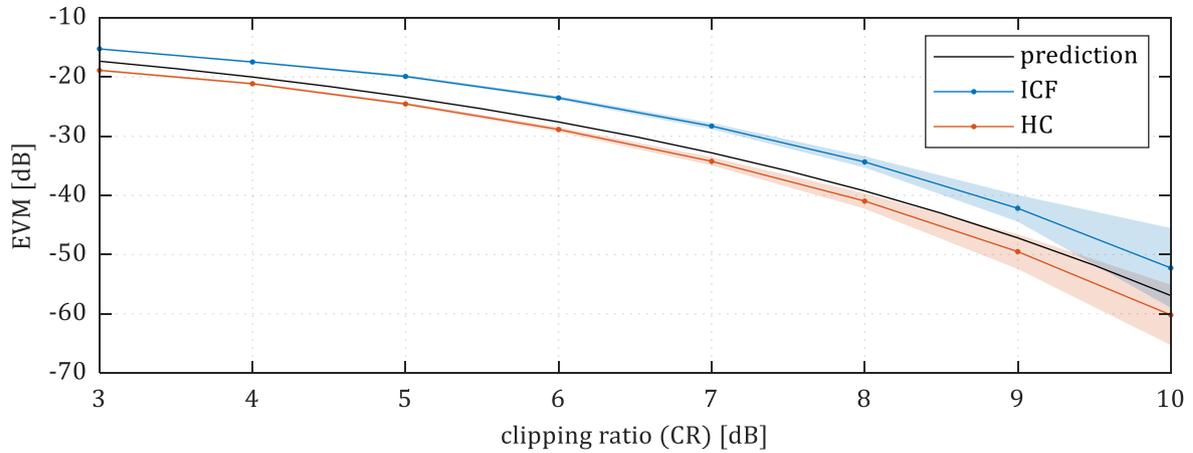


Figure 4.17: Prediction of per antenna EVM compared with the hard clipping (HC) and the iterative clipping and filtering (ICF) simulation results.

It is visible in Figure 4.17 that the standard deviation of the EVM increases with increasing the CR, for both the ICF and the HC. However, it is worth noting that for the EVM measured in percent, instead of decibel, the standard deviation was relatively stable over a wide range of CRs. For the ICF, it was between 0.2 and 0.3 percentage points in the range from 0 dB up to 8 dB CR, and below 0.2 percentage points for CRs of 9 and 10 dB. This suggests that the variation of the EVM performance over slots, which is out of scope of the proposed model, may be noticeable for relatively high CRs.

Based on the detailed results of the RMS EVM for each of the CR measurement points, presented in Table 4.3, it is visible that the difference between EVMs obtained with the ICF and HC was in the range from about 3.6 dB to about 7.9 dB. This shows that, for a given CR operating point, the impact of the ICF on the in-band signal quality is noticeably higher than in the case of the HC. This is in line with the expectations, stated in Section 4.4.1. It is assumed that this is caused primarily by the limitation of the OOB emissions in the ICF, leading to more distortions falling in-band in the following iterations of clipping.

Table 4.3: The relation between the clipping ratio (CR) and the error vector magnitude (EVM).

CR [dB]	3	4	5	6	7	8	9	10
EVM prediction [dB]	-17.3	-20.0	-23.4	-27.6	-32.8	-39.2	-47.2	-56.9
ICF's EVM [dB]	-15.2	-17.5	-19.9	-23.5	-28.3	-34.3	-42.2	-52.3
HC's EVM [dB]	-18.9	-21.1	-24.6	-28.9	-34.2	-41.0	-49.5	-60.2

Fitting the EVM prediction model to the simulation results (in the range of CRs visualized in Figure 4.17) leads to an adjusted R-squared of 0.9980 for the ICF and 0.9996 for the HC. The maximum deviation of the fitted prediction model was about 0.9 dB for the ICF and about 0.6 dB for the HC. This shows that a clipped complex Gaussian signal could be potentially used to relatively accurately predict the per antenna EVM, provided that parameters of the model are aligned with the applied peak power limitation method.

The accuracy obtained with the fitted model suggests that the power of the clipping distortion of a complex Gaussian signal could be used to predict the EVM resulting from PAPR reduction resulting either from the ICF or a linearized PA (modeled with the HC). This proves that the proposed per antenna error model could be used as a basis for prediction of the per user error.

Since the EVM results were already obtained for both the ICF and HC, for a wide range of CRs, they are used directly in the analysis of the simulation results covering the per user EVM, in the following sections. The exploration of the validity of the per user error models starts with establishing the performance in best- and worst-cases of UE angular locations, covered in Section 4.7.2.

4.7.2. Best- and Worst-Cases for Line-of-Sight Channels

Using the per antenna EVM results from the previous section as a basis, the analysis of the per user EVM starts with consideration of the performance in best- and worst-case (introduced in Section 4.4.2.3) angular locations of UEs in LoS channels.

This kind of behavior dependent on the angular location is expected for nonlinear distortion, as explained in Section 4.3.3, but not for uncorrelated noise. Hence verification of the quantization noise model was not considered from that perspective.

Definition of spatial scenarios.

Examples of best- and worst-case spatial scenarios for ULA using 32 antennas are provided in Table 4.4. These best-case angular locations were identified based on Monte Carlo numerical simulations of the expected impact of the IMPs, using the precoding-based model. These simulations covered one million scenarios of random UE placement, within the assumed 120 deg sector covered by the antenna array.

Assuming deterministic free-space LoS channels, the expected powers of the received distortion were calculated based on the radiation patterns of IMPs (predicted based on precoding, without performing any actual data transmission). The selected best-case scenarios had the minimum sum power of received IMPs. The worst-cases were selected

based on equal differences of the linear phase gradients between UEs of $2\pi/K$ radian, in alignment with the rule introduced in Section 4.4.2.3, with UEs placed symmetrically in relation to the boresight direction of 0° .

Table 4.4: Definition of the best- and worst-case scenario examples for ULA with 32 antennas.

Number of UEs (K)	azimuth angles [deg]	
	best-case scenario	worst-case scenario
1	0.0	30.0
2	-54.3, -30.0	-30.0, 30.0
3	-24.9, 1.0, 19.2	-41.8, 0.0, 41.8
4	-54.3, -30.0, -14.5, -3.6	-48.6, -14.5, 14.5, 48.6
6	-38.7, -3.6, 22.0, 25.9, 38.7, 48.6	-56.4, -30.0, -9.6, 9.6, 30.0, 56.4

Note that the scenarios for 1 UE are defined here for completeness, to confirm the expectation that there is no visible dependency of the EVM on the steering angle in this case.

Performance in free-space LoS.

First, the performance of the ICF in best- and worst-case scenarios was verified for transmission of 100 slots in free-space LoS channels. The CR was configured on the 6 dB level. This operating point was selected since the resulting per antenna EVM of -23.5 dB for ICF (obtained from the simulations described in Section 4.7.1) falls between the 3GPP EVM limits, cited in Section 4.3.1, for 64QAM (i.e., 8.0% or about -22 dB) and 256QAM (i.e., 3.5% or about -29 dB).

Based on the per slot EVM results obtained as RMS of per UE EVMs, the statistics presented in Table 4.5 were prepared. Note that in most of the cases the variation of EVMs between UEs was negligibly low (the highest variation from the RMS value per slot was up to about 0.2 dB, observed in the best-case scenario with 6 UEs), hence the RMS taken over UEs describes the overall performance relatively well.

The RMS EVM results presented in Table 4.5 clearly show that the difference in the EVM between best- and worst-cases increases with increasing the number of UEs, as expected based on the reasoning provided in Section 4.4. Also, the RMS EVM calculated for the worst-case scenarios was roughly constant irrespective of the number of used UEs. The largest deviation from the -23.5 dB level (expected based on the EVM obtained per antenna for CR of 6 dB in the results presented in Section 4.7.1) was 0.6 dB, observed in the case of 4 UEs. That largest observed difference was less than 2 standard deviations away from the expected EVM for the ICF model, based on the data presented in previous section in Figure 4.17 (i.e., 0.33 dB standard deviation of EVM for the CR operating point of 6 dB).

Table 4.5: RMS of EVM (RMS over UEs per slot) over 100 slots for free-space LoS.

Number of UEs (K)	RMS of per slot EVM [dB]				
	1	2	3	4	6
best-case	-23.5	-25.2	-26.9	-27.9	-29.0
worst-case	-23.5	-23.6	-23.8	-24.1	-23.1
Δ (best – worst)	0.0	-1.6	-3.1	-3.8	-5.9

At the same time, the differences of standard deviations between the best- and worst-cases were relatively small. Up to 0.02 dB difference was observed. This suggests that the variation of the EVM results was primarily related with the differences of the wanted signals, not the angular locations of UEs.

Comparison of the LLS results with the prediction.

A comparison of the results obtained from simulations, summarized in Table 4.5, with the prediction based on the expectations from the precoding-based model is presented in Figure 4.18, with solid lines showing RMS and filled ranges representing the distances between the RMS and 1 standard deviation of EVM over the simulated slots. The simulation results demonstrate a relatively good alignment with the prediction.

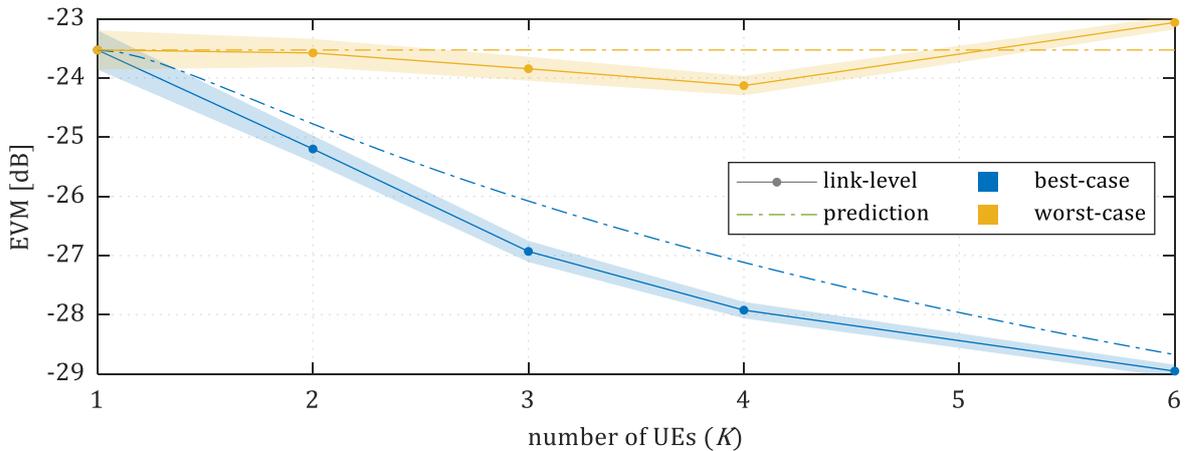


Figure 4.18: EVM statistics (RMS: solid; RMS \pm standard deviation: filled ranges) for the best- and worst-cases of UE locations in free-space LoS channels.

The RMS EVM for the best-case was lower than the prediction based on the best-case IM3s by up to about 0.9 dB. Part of that difference may be attributed to the limited simulation length. However, it is expected that this observed difference may primarily come from higher-order IMPs, especially that the impact of the fifth-order IMPs was considered when selecting the best-case scenarios (for up to 4 UEs). Reasoning for this assumption is provided in Appendix Section D.1.

The obtained results suggest that the proposed model of per user error (either the precoding-based or the best-case combined with identification of unfavorable combinations

of UE locations) could be used to predict the EVM, for transmission in free-space LoS channels, provided that the observed accuracy is considered to be acceptable.

The LLS results confirm the following expectations for the performance in deterministic LoS channels:

1. The power of the user-directed IMPs reduces with increasing the number of UEs.
2. There are combinations of angular directions that lead to a worst-case performance, independent from the number of UEs, caused by all the spatially filtered IMPs radiating towards the users.

Performance in multi-path channels with LoS.

In practice, outside of the controlled environment of an anechoic chamber, spatially correlated multi-path channels with frequency-selective fading are expected. Therefore, considering the free-space LoS results as reference, multi-path scenarios with CDL-D (LoS) channel model, using 100 ns RMS delay spread were simulated for 100 slots, for the same angular locations of UEs as in the case of deterministic LoS channels.

The comparison of the EVM simulation results, visualized in Figure 4.19, confirms that a considerable difference between the EVMs for the best- and worst-case angular locations of UEs can be expected for transmission in multi-path LoS channels, similarly like for single-path LoS.

The results for CDL-D in the best-case scenarios were closely aligned with deterministic LoS channel for up to 3 UEs. However, for 4 and 6 UEs, the EVM results with multi-path LoS channels were higher than for single-path LoS. It is expected that this was caused by two factors. First, the best-case angular locations calculated for free-space LoS channels do not take into account the multitude of angles of the scattering clusters in CDL-D channels. Hence, not only IMPs radiate in more directions than for free-space LoS, but also the signal is received by the UEs from many more angles. This leads to higher impact of the spatially filtered IMPs, compared to transmission in a deterministic LoS channel. The second factor is the variation of the per antenna power, leading to a higher sum power of the introduced IMPs. That aspect is explored in Appendix Section D.2, together with the performance in NLoS channels.

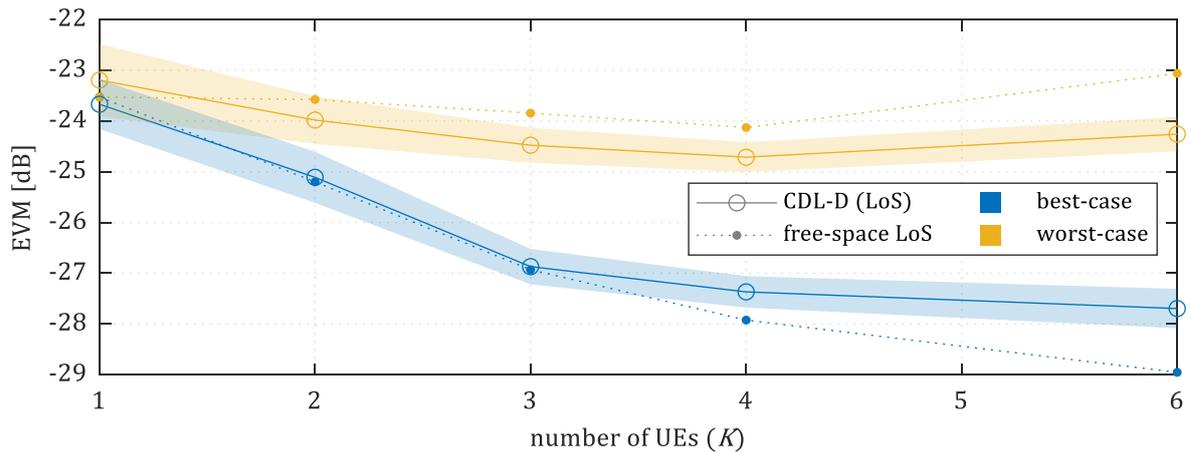


Figure 4.19: Comparison of EVM simulation results for multi-path with LoS (RMS: solid; RMS \pm std dev: filled ranges) and free-space LoS, for UEs located at the sets of angles of the best- and worst-case scenarios.

Note that the results shown in Figure 4.19 represent a power allocation strategy without equal power allocation over SCs. The same verification was performed for equal distribution of power over SCs. This led to about 0.2 dB improvement of the RMS EVM results for both best- and worst-cases, which suggests that the power budget with equal power distribution over SCs may be marginally better from the perspective of the impact of residual TX impairments, at least for multi-path LoS transmission. This may be expected since that approach equalizes the impact of distortion over SCs and between UEs.

The results for best- and worst-case angular locations of UEs, presented in Figure 4.19, suggest that for channels with a significant LoS component the EVM resulting from the peak power limitation can noticeably depend on the angular locations of UEs. This is in line with the expectations given in Section 4.4.2.

It is also clearly visible that there are worst-case sets of angular locations in LoS transmission, both for free-space and multi-path propagation, which have considerably worse EVM. This suggests that either a scheduling algorithm could avoid spatially multiplexing UEs positioned at angles that lead to the identifiable worst-cases or link adaptation could select appropriate modulation orders and code rates, in anticipation of a degraded EVM performance.

4.7.3. Statistics of the Error Level for Randomly Placed Users

The EVM results for best- and worst-cases of angular locations, described in Section 4.7.2, provide important boundary conditions for forming expectations regarding the EVM that can be achieved for randomly placed users.

First, a comparison of the EVM performance predictions from the best-case model (introduced in Section 4.4.2.1) and the precoding-based model (proposed in Section 4.4.2.4), is explored in Section 4.7.3.1. Next, the obtained predictions are compared against the EVM performance results from LLS for ICF, in Section 4.7.3.2, to verify the prediction accuracy. Finally, the impact of the array geometry is explored in Section 4.7.3.3.

Apart from judging the accuracy of the proposed EVM prediction models, the intention of this verification is to establish an understanding of the typical EVM performance, assuming random placement of users. The focus is especially on clarifying how far the typical EVM performance can be from the best-case EVM prediction by (4.19), depending on the number of spatially multiplexed UEs K , number of antennas M , and the array geometry. The results of this study are meant to inform the selection of the approach for modeling of the impact of distortion on the achievable DL data rates, handled in Chapter 5.

4.7.3.1 Prediction from Numerical Simulation for Linear Arrays

First, the results of the numerical simulations (Monte Carlo trials with varying UE locations in free-space LoS channels) for the precoding-based model were analyzed to understand how likely different EVM levels are in LoS scenarios. Statistics of the power of spatially filtered IMPs received by UEs, in relation to all IMPs radiated from a base station using either 32 or 128 antennas M , are visualized in Figure 4.20 for a range from 2 to 8 spatially multiplexed UEs K . The results show that the portion of the received power of the spatially filtered IMPs is consistently much smaller for the array with more antennas, which is in line with the expectation that increase in the array aperture can positively influence the results by providing finer angular resolution. This confirms the expectation that the statistics of EVM are influenced not only by the number of UEs, but also by the number of antennas.

It can be observed in Figure 4.20 that typically (i.e., at the CCDF level of 0.5) the portion of the received spatially filtered IMPs increases when the number of UEs is increased. For the number of UEs K from 2 to 6, less than 10% of cases are expected to receive more than a quarter of the power of radiated spatially filtered IMPs in case of $M = 32$. These results suggest that for transmission with at least 32 antennas, typically the UEs in LoS channels are expected to receive only a minor portion of the radiated spatially filtered IMPs.

At the same time, for 6 and 8 UE cases for 32 antennas almost all scenarios of angular location of UEs are expected to lead to higher than -10 dB portion of received IMPs. While for 128 antennas, visibly lower portion of the power of spatially filtered IMPs can be

expected to be received. This suggests that a minor EVM performance degradation, resulting from spatially filtered IMPs, could be expected when the M/K ratio reduces.

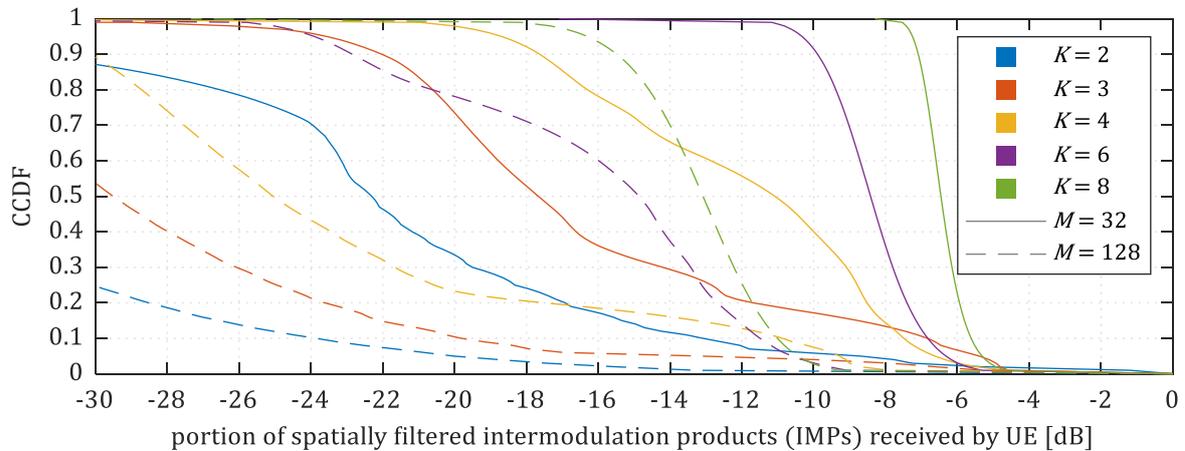


Figure 4.20: Prediction of the portion of spatially filtered intermodulation products (IMPs) received by UEs, based on numerical simulations of $1e6$ Monte Carlo trials.

The statistics of the portion of the power of spatially filtered IMPs that is expected to be received by UEs, presented in Figure 4.20, can be used to reflect the impact of both the UE angular locations and angular granularity of the array. In the best-case, a negligible portion of spatially filtered IMPs is received, leading to EVM performance estimate based on the user-directed IMPs reflected in the prediction model defined in (4.19). In practice, depending on the UE angular locations and the spatial granularity of the array, a portion of the spatially filtered IMPs power may visibly increase the total power of the received distortion.

Such a combination of the best-case EVM performance, limited by the received power of the explicitly user-directed IMPs, with the statistics of the received power of spatially filtered IMPs is presented in Figure 4.21 for a per antenna EVM of -23.5 dB. Clearly, the EVM performance could be expected to fall in a range of EVMs strongly depending on the number of UEs K , in alignment with the general expectation from (4.18).

A significant impact of the number of antennas M on the EVM performance is also clearly visible in Figure 4.21, especially when the number of UEs was increased. This is in line with the expectation of the influence the angular granularity (determined by the array aperture) has on the reception of the spatially filtered IMPs, as considered in Section 4.4.2.4. As visible in Figure 4.21, the expectation of the EVM performance obtained based on the precoding-based model from Section 4.4.2.4, a gap between the typical EVM performance (i.e., at 0.5 CCDF level) and both best- and worst-case EVM performance increases with increasing number of UEs.

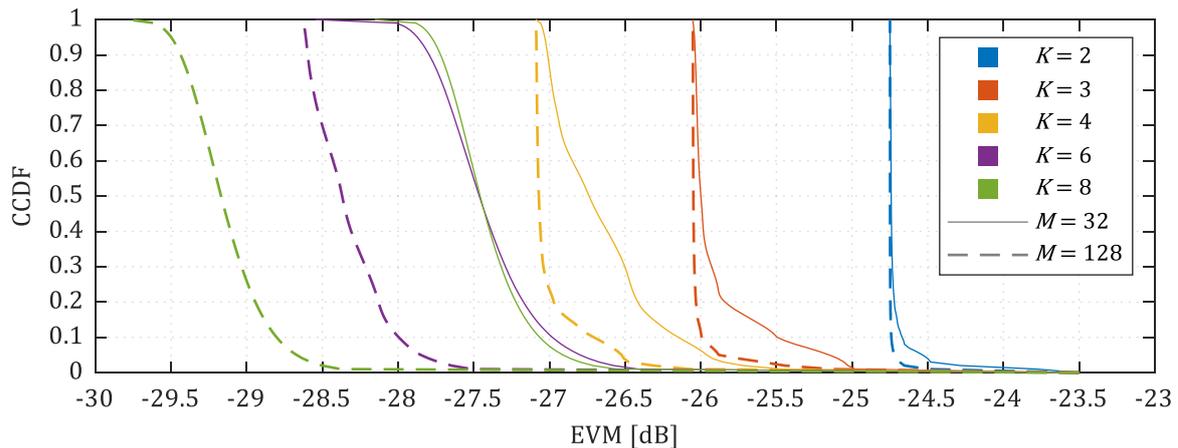


Figure 4.21: Expected EVM distribution, over spatial scenarios, for a free-space LoS channel for different number of UEs K for 32 (solid lines) and 128 (dashed lines) antennas M . Based on the numerical simulations of third-order IMPs, assuming -23.5 dB EVM for $K = 1$.

A typical EVM improvement (i.e., the difference between the -23.5 dB EVM level, experienced generally for $K = 1$ and in the worst-case with any K , and the 0.5 level of the EVM CCDF) observed for the precoding-based model from (4.38) is compared in Table 4.6 with the improvement observable with the simplified best-case model from (4.19). The largest deviations between the models are visible for high K and low M , as expected. This suggests that the best-case model cannot be generally used standalone, since it may lead to noticeably overoptimistic estimates. Extension of the best-case model with identification of the unfavorable combinations of the UE angles could likely improve the prediction accuracy.

Table 4.6: Predicted improvement of the per user EVM in relation to the per antenna EVM.

Number of UEs (K)		Improvement in EVM [dB]					
		1	2	3	4	6	8
$M = 32$	best-case model	N/A	1.3	2.6	3.6	5.2	6.0
	precoding-based model	N/A	1.2	2.5	3.2	4.0	4.0
$M = 128$	best-case model	N/A	1.3	2.6	3.6	5.2	6.3
	precoding-based model	N/A	1.2	2.5	3.6	4.9	5.7

Next, the EVM performance for ICF measured in LLS is presented in Section 4.7.3.2, considering also the EVM performance prediction accuracy of the proposed models.

4.7.3.2 Link-Level Simulation for Linear Arrays

An initial verification of the EVM in LLS for randomly placed UEs was presented in [62] (co-authored by the PhD candidate) for HC. Additional insights from these simulation results are given in Appendix Section D.3. Here, the attention is given to the ICF.

First, the performance in free-space LoS scenarios was verified, for $M = 32$. The EVM results obtained from LLS (with 10 slots per spatial scenario) for 100 scenarios of UE placement, for a range of users from 1 to 8, were compared with the prediction of the EVM from the numerical simulations of precoding-based model covered in Section 4.7.3.1 (representing one million scenarios of UE placement).

The comparison of the CCDFs of EVM, representing the distribution of the RMS of per UE EVMs over spatial scenarios, is presented in Figure 4.22. At first glance, it seems that the EVM prediction is typically pessimistic, which is in line with the expectation when using a model based on third-order IMPs to predict the EVM performance of a nonlinear system which introduced also higher-order IMPs. This is especially visible for 3 and 4 UEs, where about 0.6–0.7 dB gap between prediction and link-level can be observed in the simulation results. Additionally, it is not excluded that a part of that apparent gap in the EVM performance between prediction and link-level can be attributed to the moderate number of 100 spatial scenarios verified in LLS.

It is clear that usage of the precoding-based prediction model provides more accurate estimates of the EVM performance than the simplified best-case prediction model. This is especially visible for 6 and 8 UEs, where about 5.2 dB and 6.0 dB improvement could be expected respectively in the best-case per Table 4.6, while the CCDFs of EVM in Figure 4.22 obtained from the link-level results reveal that a typical EVM improvement was noticeably lower (and closer to the results of the precoding-based prediction).

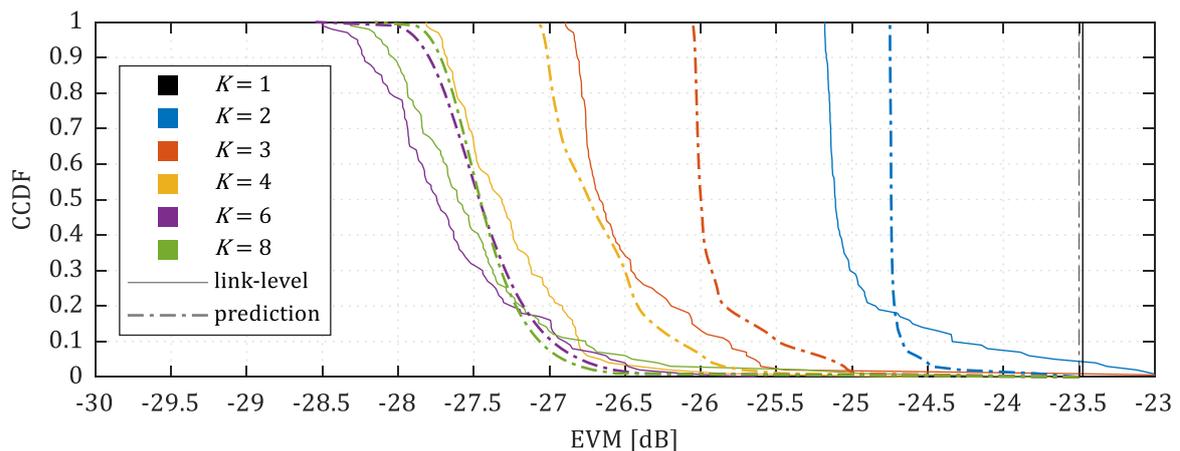


Figure 4.22: Comparison for the statistics of EVM (RMS of per UE EVMs), over spatial scenarios, of link-level simulation results (solid lines) with the prediction based on the numerical simulations using precoding weights (dash-dotted lines), for a free-space LoS channel for different number of UEs K for 32 antennas.

Note that there was a handful of 2 and 3 UE scenarios that led to higher EVM than for 1 UE, noticeable in the tail of the CCDF in Figure 4.22. Analysis of the related outlier cases

confirmed a somewhat worse performance of the ZF precoding (resulting in reduced power of the received wanted signal) in these cases, which may explain the EVM degradation.

Accuracy of the prediction.

Comparison of the statistics of EVM performance for link-level and prediction was done, by calculating the differences between the percentiles of EVM, in order to judge the accuracy of the prediction. Results of that comparison are presented for different levels of percentiles of EVM for free-space LoS channels (in alignment with the CCDFs presented in Figure 4.22) in Figure 4.23(a) and for multi-path LoS channels (CDL-D) in Figure 4.23(b). Note that the EVM values in Figure 4.23 are increasing with the increasing percentile levels, which is aligned with the way a cumulative density function (CDF) is defined (which is reflected in reverse in CCDF, with the higher EVM values in Figure 4.22 related with the lower CCDF values).

It is generally visible in Figure 4.23 that the deviation between the link-level and prediction was lower than 1 dB, for both free-space and multi-path LoS results. The bias towards pessimistic prediction (i.e., negative value of the EVM difference) is clearly visible in most of the cases (for percentiles up to the 50th, as well as any percentile for 3 and 4 UEs). A visibly higher deviation for CDL-D compared with free-space LoS can be observed for 2 UEs (a lower EVM in the CDL-D results, across all percentile levels) and 1 UE (more deviation generally for CDL-D than for the free-space LoS, apart from the 50th percentile).

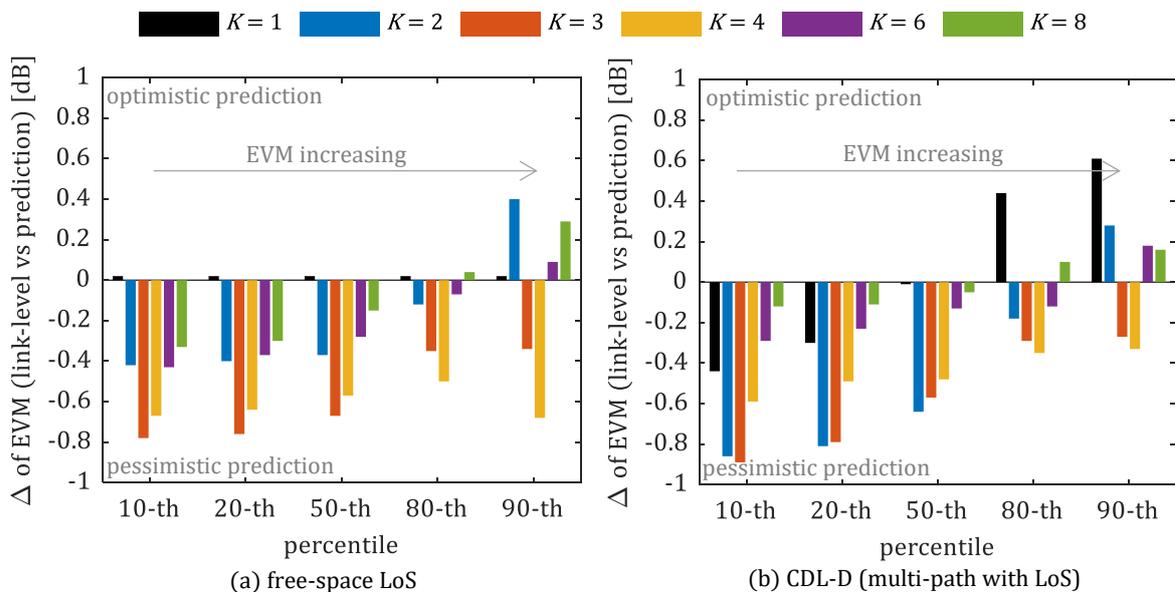


Figure 4.23: The differences of percentiles of EVM between the link-level simulations and precoding-based prediction for different number of UEs (K) for transmission using 32 antennas in deterministic LoS and frequency-selective channels with LoS.

Overall, it seems that usage of a model basing solely on third-order IMPs to predict the EVM performance of ICF (which inherently produces higher-order IMPs) has led to generally pessimistic prediction, with a typical deviation (i.e., at the 50th percentile) of up to about 0.7 dB. It is important to emphasize that this was true not only for the free-space LoS scenarios, which are aligned with the way the model based on beam steering was defined, but also for the multi-path LoS scenarios.

It is assumed that the accuracy of the predictive model could be improved, at least in the free-space LoS case, by employing the higher-order IMPs in the prediction. However, potentially a different viable option would be to reduce the inherent pessimistic bias of the third-order IMPs based model, by providing a correction depending on the number of UEs.

While the proposed EVM prediction models are defined based on a free-space LoS assumptions, the feasibility of using them generally to cover also multi-path scenarios with LoS was demonstrated here.

4.7.3.3 Impacts of Spatial Resolution and Elevation Steering for Rectangular Arrays

Basing on the apparent viability of using EVM estimates obtained based on precoding weights, using numerical simulations based on Monte Carlo trials with varying UE locations, for predicting the EVM performance in LoS transmission from ULA, covered in Section 4.7.3, a similar approach was taken for URA. The angular ranges for UE locations were limited to $\pm 60^\circ$ in azimuth and $\pm 20^\circ$ in elevation.

Verification of the validity of the prediction model.

First, the feasibility of using the precoding-based prediction model was verified using a square URA with 64 antennas. The LLS results (RMS EVM over 10 slots) based on 50 scenarios of randomly placed UEs were compared with the numerical simulations of one million Monte Carlo trials of UE locations. Note that in this case after verification of the results against the 1 dB rejection threshold, a substantial number of results were ignored for $K > 4$ (about 30% for $K = 6$ and about 50% for $K = 8$, in both link-level and prediction cases). The comparison of CCDFs for link-level and prediction, visualized in Figure 4.24, demonstrates that a similar prediction accuracy as in case of ULA could be expected. This suggests that the precoding-based model could be expected to relatively well reflect the EVM statistics for rectangular arrays.

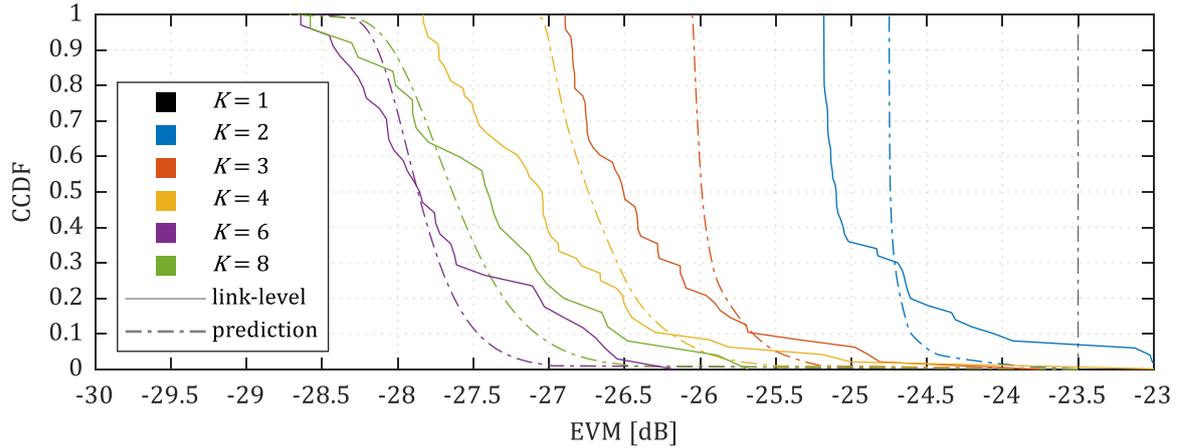


Figure 4.24: Comparison of the statistics over spatial scenarios of RMS of per UE EVMs, for link-level simulation results (solid lines) with the numerical simulations for the prediction based on precoding (dash-dotted lines), for a free-space LoS channel for different number of UEs K for a square URA with 64 antennas.

Separate limitation of the azimuth and elevation granularity.

Using the expectations of the EVM generated based on numerical simulations with the precoding-based model, the relevance of the array geometry, not only the number of antennas M , is demonstrated in Figure 4.25 by comparing the EVM statistics for ULA with 32 antennas and URA with the same number of 32 columns but a total of 256 antennas (i.e., 8 rows). When the UEs are allowed to be placed at random azimuth locations but constrained to the same elevation angle, which was 0° here, the rows of URA do not visibly improve the results obtained with ULA. This makes sense, since usage of equal elevation angles for all the UEs is expected to lead to generation of IMPs at the same elevation angle.

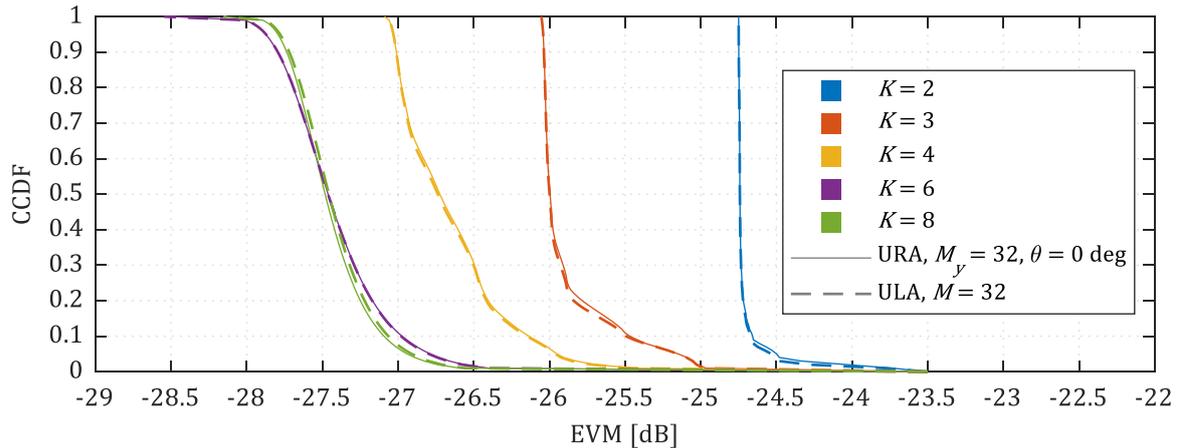


Figure 4.25: In the worst case of equal elevation steering (here $\theta = 0^\circ$) for all UEs, the EVM statistics for URA depend only on the number of columns M_y , not on the total number of antennas M (here 256 for URA). Based on the EVM prediction from the numerical simulations of IMPs, assuming -23.5 dB EVM for $K = 1$.

Dependence on the angular resolution.

In the general case, when the elevation steering is not artificially constrained to the same angle for all UEs, the spatially filtered IMPs are generally expected to radiate to elevation directions that are different than UE elevation angles, with certain classes of scenarios leading to at least partial radiation towards the UE locations, as explained in Section 4.4.2.2, with all the spatially filtered IMPs radiating towards UEs in the worst-case, as clarified in Section 4.4.2.3. In order to capture that spectrum of spatial scenarios, statistics of EVM based on prediction from numerical simulations were generated for UEs uniformly randomly placed between -20° and $+20^\circ$ elevation. The comparison of EVM statistics for URAs with 16 and 32 columns M_y (with 8 rows in both cases) is presented in Figure 4.26, based on numerical simulations for ZF precoding weights generated for transmission in free-space LoS channels to K UEs. Better results obtained with $M_y = 32$ are expected to be resulting from a finer angular granularity.

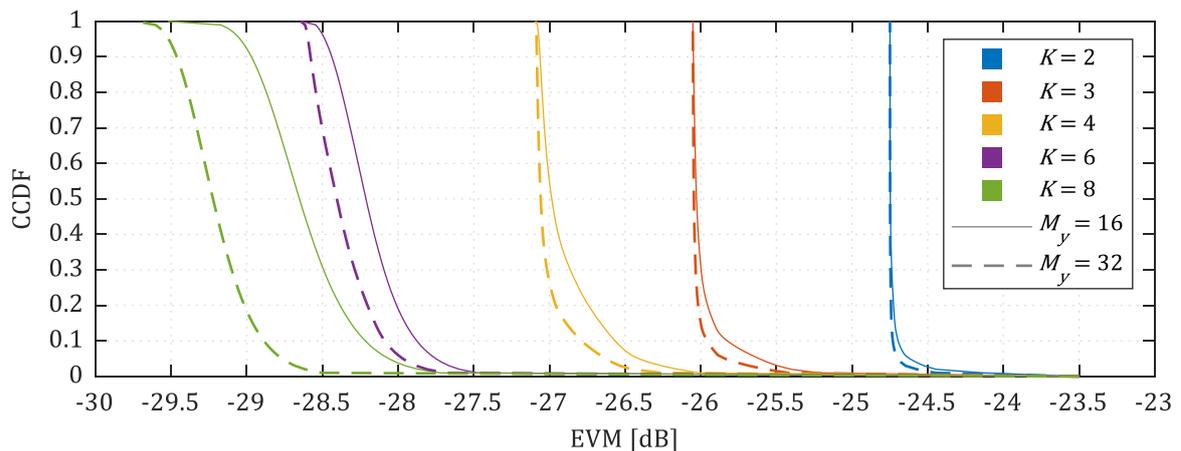


Figure 4.26: Comparison of the expected EVM distribution for URAs with 16 and 32 columns M_y (both with 8 rows). Based on the numerical simulations of predicted 3rd-order IMPs, assuming -23.5 dB EVM for $K = 1$.

When comparing the results of $K > 2$ for scenarios using equal elevation steering in Figure 4.25 and randomly selected elevation in Figure 4.26, it is not only visible that URA using 32 columns led to better performance than ULA with 32 antennas when the constraint on the elevation angle was lifted, but also the results for URA with 16 columns were better than for ULA. That improvement was especially visible for the results for 6 and 8 UEs. Actually, the results obtained with URA using 256 antennas are comparable to EVM performance for ULA with 128 antennas, visualized in Figure 4.21.

4.7.4. Performance in Link-Level Simulation for Non-Line-of-Sight Channels

The EVM performance prediction models proposed in Section 4.4.2 are dedicated for LoS scenarios. Hence, the performance in NLoS scenarios is characterized here solely based on LLS results, without an a priori prediction of the EVM performance. This consideration of 100 scenarios (with different angular placements of UEs) is complemented by the EVM results obtained for NLoS for the best- and worst-case sets of UE locations (identified for LoS), covered in Appendix Section D.2.

The CCDFs of link-level EVM results, representing the distribution over spatial scenarios of the RMS of per UE EVMs, for transmission from ULA with 32 omni-directional antennas over CDL-A NLoS channels are presented in Figure 4.27 for a range of spatially multiplexed UEs from 1 to 8. Clearly, the characteristic for 1 UE seems to be an outlier when compared to the results for $K > 1$, with the typical EVM level about 3 dB higher for $K = 1$. The CCDFs of EVM for $K > 2$ seem to have comparable characteristics over different percentile levels, with up to about 0.5 dB distance between the CCDF curves. This suggests that the case of 1 UE may need to be considered as a special case. Some more insights into the performance with single user in NLoS channel are provided in Appendix Section D.2.

While for 2 UEs the 50th percentile level virtually matched the EVM performance obtained with 8 UEs, the spread of EVMs between minimum and maximum values of the CCDF observed in the results for $K = 2$ was relatively large compared to the other cases (it was about 4.4 dB, which was even larger than about 3.9 dB range observed for 1 UE).

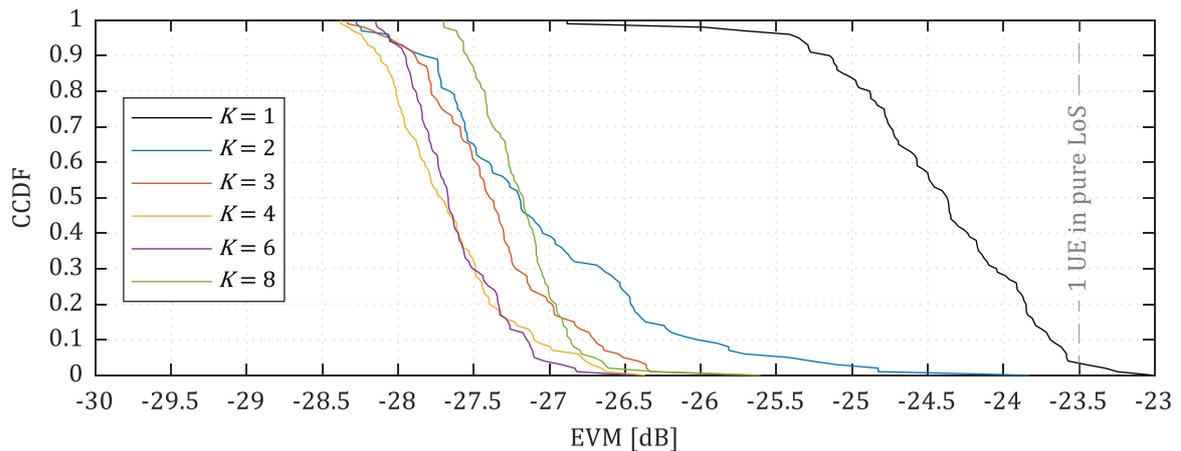


Figure 4.27: Statistics of EVM (RMS of per UE EVMs, covering transmission of 10 slots) over 100 spatial scenarios for link-level simulation results for CDL-A (NLoS) channels for different number of UEs K for 32 antennas.

It is expected that when the per antenna power is controlled in a way that the variation of power over antennas is decreased, a reduction of the dependence of the EVM performance

on the number of UEs could be observed for NLoS, similarly as observed in the verification with equal CR over antennas covered in Section D.2.

The LLS results obtained for NLoS show that for $K > 1$ similar EVM levels, irrespective of the number of UEs could be expected for the case of 32 antennas. This suggests that reflection of angular positions of UEs in the EVM prediction may be significantly less important for NLoS than for LoS cases. However, a dependence on the number of antennas (restricting the angular granularity of radiation) may still need to be considered when predicting the EVM performance in NLoS scenarios, similarly to the LoS cases covered by the proposed prediction models.

The validity of that expectation was separately confirmed for 2 random scenarios with CDL-A for $K = 6$ in which about 1.6 dB improvement in EVM was obtained when increasing the number of antennas from 32 to 128. In fact, even for the 1 UE results shared in Figure D.3 some EVM improvement when going to 128 antennas can be observed. However, in the end the cases with $K = 1$ may need to be treated as outliers, especially if the power variation over antennas is not explicitly controlled.

4.7.5. Impact of the Power Allocation Strategy

So far, the impact of the residual TX impairments on the EVM performance was verified here for the strategy of equal power allocation between UEs. The LLS results confirmed validity of the expectation that the error level in transmission over LoS channels depends on the number of UEs. It was demonstrated that increasing the number transmission layers leads generally to reduction in EVM for nonlinear distortion, as shown in Section 4.7.3. While it leads to increase in EVM for uncorrelated noise, as explained in Section 4.5 and separately demonstrated in Appendix Section D.4.

The effects of distortion should also be considered in the alternative power allocation strategy, with the power distributed between UEs inversely proportional to their path losses (targeting comparable SNRs of UEs), in order to evaluate the fairness of this power allocation approach which aims at providing roughly equal QoS to all UEs.

The behavior of distortion for the case of unequal power allocation between UEs can be exemplified based on the 4 UE scenario defined in [25] (with distances from the base station of 50, 75, 100, and 500 m). Here, free-space LoS channels are assumed for each of the UEs. The resulting differences between path losses lead to a highly directional radiation pattern, with most of the power allocated to the UE with the highest path loss. This is visible in the

comparison of the radiation patterns for the wanted signal x between the two power allocation strategies, visualized in Figure 4.28 for ULA with 32 omni-directional antennas.

Uncorrelated Gaussian noise.

Clearly, when the distortion signal d is radiated uniformly over the azimuth angles (as expected from the impairments modeled by uncorrelated Gaussian noise) the SDRs for the equal power case could be expected to be roughly equal (leading to similar EVM performance over UEs). However, the SDRs in the equal SNR case could be expected to be visibly different (resulting in the expectation of different EVM performance between UEs).

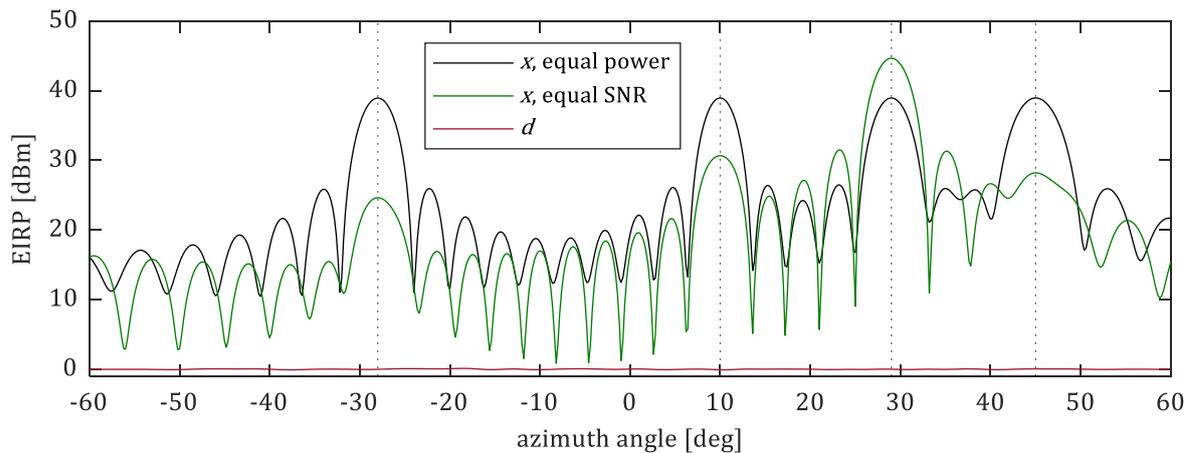


Figure 4.28: Comparison of the radiation patterns of the wanted signal x for equal power and equal SNR allocation strategies, for linear array with 32 antennas. Uniformly radiating distortion d shown for reference.

A comparison of the per UE EVM obtained for the two considered power allocation strategies (i.e., the equal transmit power and the transmit power inversely proportional to the path loss, targeting equal SNR) in LLS with the quantization noise model, is given in Table 4.7. These results are similar to the results shared originally in [62] (co-authored by the PhD candidate). While the power allocation in case of the comparable QoS approach was modified here to obtain exactly same SNR for all the UEs, the same conclusion is reached.

Based on this example, it is visible that the cost of using the equal SNR strategy is not only the apparent SNR reduction for some of the UEs (modified by the change in the transmit power) but also degradation of the EVM performance. In this case, a notable increase in the EVM for 3 out of 4 UEs (these closer to the base station) can be observed in Table 4.7. At the same time, an improvement of about 5.7 dB in both the EVM and the SNR was obtained for the UE with the highest path loss. This may result in failing to achieve the objective of comparable QoS for all the spatially multiplexed UEs.

While the allocation of power to compensate for the path loss differences led to SNRs on the same level, it is possible that not only the sum data rate but also the per user data rates

with this kind of approach would be reduced compared with the equal power option, depending on the actual EVM and SNR levels. This is explored later, in Section 5.3.2.

Table 4.7: Examples of per UE transmit powers and EVMs, for two different power allocation strategies: equal power; power distribution based on path loss, to achieve equal SNR.

UE number (k)	distance [m]	azimuth angle [deg]	path loss [dB]	equal power strategy		equal SNR strategy	
				allocated TX power [dBm]	EVM [dB]	allocated TX power [dBm]	EVM [dB]
1	50	-28	77.3	24.0	-38.1	9.7	-23.8
2	75	45	80.8	24.0	-38.0	13.2	-27.3
3	100	10	83.3	24.0	-38.1	15.7	-29.8
4	500	29	97.3	24.0	-38.1	29.7	-43.8

It is worth to point out that in case of an uncorrelated noise, the apparent EVM performance degradation in the equal SNR strategy (experienced by the UEs closer to the base station) could be theoretically circumvented by increasing the number of antennas. However, for the nonlinear distortion, characterized generally by directional radiation patterns, the increase in the number of antennas may not necessarily compensate fully for the EVM performance degradation. Hence the EVM performance with ICF in the same 4 UE scenario for both power allocation strategies is characterized next.

Nonlinear distortion.

Clearly the radiation patterns of distortion d for the ICF, visualized in Figure 4.29, differ fundamentally from the uniform radiation patterns of distortion obtained in case of the quantization noise model. The comparison of distortion patterns for the different power allocation strategies (i.e., equal power and equal SNR), obtained here as a result of LLS (with ICF as the only notable source of distortion), reveals different distribution of distortion power over azimuth angles.

In case of the equal power strategy, the highest peaks were aligned with the UE directions and the radiation pattern varied in about 10 dB range over the visualized angles. While in case of the equal SNR strategy, the variation of the distortion power over angles was visibly increased, to some extent in alignment with the unequal power allocation reflected in the radiation pattern of the wanted signal. However, it seems clear that the SDRs of the UEs with reduced power (i.e., UE#1 at -28° , UE#2 at 45° , and UE#3 at 10°) were negatively impacted, since the reduction of the distortion power d in their directions was visibly smaller than the reduction of the power of the wanted signal x .

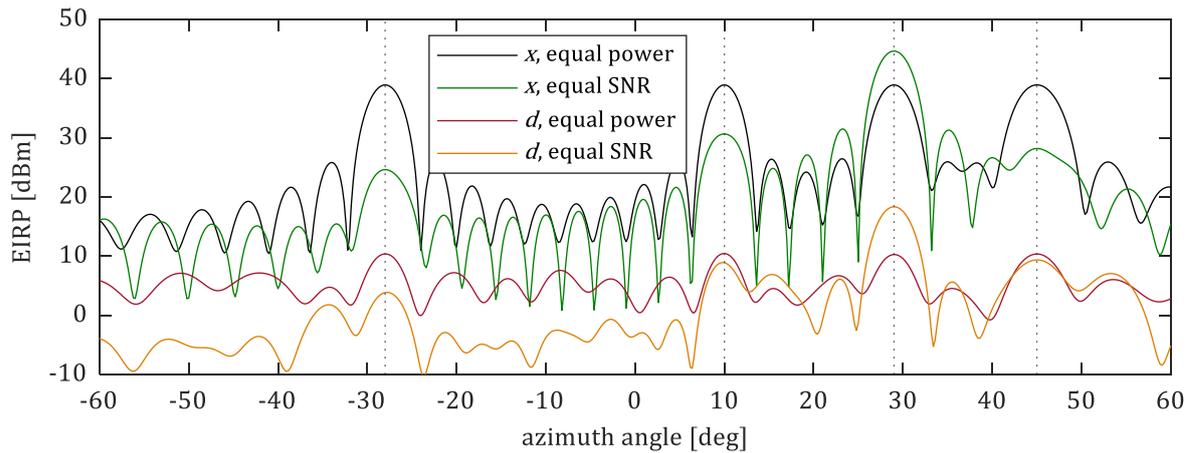


Figure 4.29: Comparison of the radiation patterns of the wanted signal x and nonlinear distortion d for equal power and equal SNR allocation strategies, for linear array with 32 antennas.

In order to obtain a quantifiable example of the impact of an unequal power allocation between UEs, the EVM performance of ICF using CR of 6 dB was characterized based on LLS. The per UE EVM results obtained for both power allocation strategies for arrays with 32, 128, and 512 antennas M are summarized in Table 4.8. These results suggest that unequal power allocation can lead to EVM performance degradation for all the UEs. It seems that when a significant portion of the total power is allocated to one UE, as in the case of equal SNR strategy for this scenario (where UE#4 uses 29.7 dBm out of the 30.0 dBm power budget), the UEs with a minor power allocation can be expected to get similar EVM levels (that are higher than for the UE with the highest power allocation), provided that the angular granularity of the array is fine enough (as demonstrated here by the apparent reduction of the variation of EVMs over the lower powered UEs when increasing the number of antennas from 32 to 128).

Table 4.8: Per user EVMs resulting from distortion introduced by ICF, for UEs in free-space LoS channels, depending on the power allocation strategy and the number antennas M .

UE number (k)	distance [m]	azimuth angle [deg]	EVM [dB]					
			equal power strategy			equal SNR strategy		
			$M = 32$	$M = 128$	$M = 512$	$M = 32$	$M = 128$	$M = 512$
1	50	-28	-27.7	-28.0	-28.0	-20.6	-21.3	-21.3
2	75	45	-27.6	-28.1	-28.1	-18.1	-21.6	-21.7
3	100	10	-27.6	-28.0	-28.0	-21.2	-21.2	-21.2
4	500	29	-27.7	-28.0	-28.0	-25.4	-25.6	-25.6

4.8. Summary and Proof of Thesis

The analytical analysis of the radiation characteristic of nonlinear distortion, covered in Section 4.4.2, revealed that the EVM performance depends on the number of UEs K , number of antennas M , and the geometry of the array (determining the angular granularity of radiation). At the same time, it was identified that a portion of the spatially filtered IMPs (normally radiating in distinct directions that are different than the angular locations of UEs) may be radiated to the direction of UEs (similarly to the explicitly user-directed IMPs). This happens when the precoding weights are selected in a way that results in equal differences between phase gradients for different transmission layers. In fact, it was revealed in Section 4.4.2.3 that when these differences in phase gradients between all the UEs are equal to $2\pi/K$ in case of a half-wavelength spacing between antennas, all the spatially filtered IMPs radiate to the directions of the UEs, forming worst-case class of scenarios characterized by a lack of EVM performance improvement compared to a SISO transmission.

Two EVM performance prediction models for nonlinear distortion, described with the peak limitation (either by saturation in the PA or by explicit PAPR reduction) as the source of distortion in mind, were proposed. Both are rooted in the expectation related with the portion of the third-order IMPs that radiate to the same directions as the precoded data signal.

The simplified best-case prediction model (based on the number of UEs K and the number of antennas M), introduced in Section 4.4.2.1, is assumed to be potentially usable for characterization of a typical EVM performance. However, addressing the issue of radiation in the UE directions of a portion of the spatially filtered IMPs (and especially addressing the worst-case combinations of phase gradients), may necessitate a dedicated solution for dealing with these outliers. The alternative approach of using a precoding-based prediction model, proposed in Section 4.4.2.4, not only addresses the issue of these EVM performance outliers but also covers the aspects related with the limited angular granularity of radiation, potentially resulting in receiving by the UE a portion of the spatially filtered IMPs radiating in distinct direction other than the UE directions (due to the relatively wide main lobes in the radiation pattern of distortion, especially in arrays with not so massive number of columns or rows of antennas).

In addition to characterization of the EVM performance for nonlinear distortion, expectations related to the EVM performance resulting from uncorrelated noise, with quantization as a source of the modeled distortion in mind, were outlined in Section 4.5. It was pointed out that while the uncorrelated noise is expected to radiate uniformly, the EVM

can be expected to degrade with increasing number of UEs (which share a common pool of TX power) when the noise scales with the power of the combined precoded signal.

The efficacy of the proposed models was verified in simulation. The EVM performance predictions obtained with the precoding-based model by the numerical simulations using Monte Carlo trials of angular locations of UEs were compared with LLS results.

First, the feasibility of using the per antenna EVM estimate introduced in Section 4.4.1 was confirmed. Utilization of the proposed model led to adjusted R-squared of 0.9980 for the ICF and 0.9996 for the HC, in the range of CRs between 3 and 10 dB. The accuracy obtained with the fitted model suggests that the proposed per antenna error model could be used as a basis for prediction of the per user error.

Next, the boundaries of the EVM performance were established by obtaining the LLS results with ICF, in both free-space LoS and multi-path with LoS, for the identified best- and worst-case sets of UE angular locations. For the free-space LoS cases, the largest observed deviation in EVM between the LLS results and prediction using the precoding-based model was about 0.6 dB for the worst-case scenarios and about 0.9 dB for the best-case scenarios (which was in line with the expectation that a prediction based on third-order IMPs should provide pessimistic estimates). In fact, usage of the simplified best-case prediction model to estimate the performance in the identified best-case scenarios would have resulted in a comparable accuracy of prediction. With the spatially correlated frequency-selective channels, it was demonstrated that for transmission with LoS (using the CDL-D channel model) a similar dependency to UE locations as in the free-space LoS could be expected. Additionally, it was revealed that the power budget with equal power distribution over SCs may be marginally better from the perspective of the impact of residual TX impairments, at least for multi-path LoS transmission.

The statistics of the EVM for randomly placed users, obtained from simulations using Monte Carlo trials with varying UE locations for both the precoding-based model (using 1 million trials) and the link-level (using either ICF or HC as the source of distortion in 100 trials), showed that the model generated generally pessimistic predictions, with a typical deviation of about 0.7 dB (for both free-space LoS and multi-path with LoS).

The feasibility of using the precoding-based model to predict the EVM performance for rectangular arrays was confirmed. It was also demonstrated that the array geometry, not only the total number of antennas, has a significant influence on the performance due to separate limitation of the angular granularity in both azimuth and elevation.

A similar approach applied to the NLoS scenarios, based on LLS results for ICF, resulted in comparable characteristics of the CCDFs of EVMs for $K > 2$ over the full range of percentile levels. This suggests a lack of strong dependence of the EVM performance on the number of UEs, which is a very different behavior than observed for LoS scenarios. While the difference between the 50th percentiles of EVM was less than 0.5 dB for the range of K from 2 to 8, the EVM performance for 1 UE was a clear outlier with about 3 dB higher 50th percentile.

The relevance of considering the relation between the power allocation strategy and the radiation characteristic of distortion, was demonstrated in LLS for both the quantization and peak power limitation models. The results suggest that usage of unequal power allocation based on the path losses experienced by UEs (to achieve similar SNRs), may generally lead to degradation of the EVM performance for the UEs that are allocated with less power than in case of the equal power allocation strategy. This suggests that the goal of providing comparable QoS to all the spatially multiplexed UEs may be not reached, despite achieving equal SNRs.

In conclusion, the analytical analysis described in Section 4.4, supported by the simulation results shared in Section 4.7, confirmed the validity of the first thesis of this dissertation. On the one hand, it was demonstrated that the EVM at the user depends on the structure of the base station. With the peak power capability of the PA influencing the per antenna EVM, the number of antennas limiting the array gain, and the array geometry impacting the beamwidths of the radiated nonlinear distortion. On the other hand, it was demonstrated that selection of the precoding weights (including the power distribution strategy) may significantly alter the EVM performance. Moreover, it was shown that in the worst-case it can be expected that distortion achieves the same array gain as the wanted data signal, radiating fully in the directions of the UEs.

Overall, the impact of the residual TX impairments on the EVM performance of Massive MIMO base station, explored in this chapter, has a potential for limiting the achievable DL data rates. Especially when targeting a high energy efficiency operating regime of the PAs (to achieve a relatively low energy consumption), leading to increasing the per antenna power of the nonlinear distortion resulting from the peak power limitation (either by saturation in the PA or explicit PAPR reduction using a method based on signal distortion). The system-level aspects of linking the EVM performance with the data rate are explored in more detail in Chapter 5.

Chapter 5

Impact of Distortion on Downlink Data Rates and Energy Consumption

In order to translate the EVM performance impact of the residual TX impairments, explored in Chapter 4, to the impact on the DL data rates and energy consumption of the base station, there are multiple assumptions that need to be made. On the one hand, both the expected EVM performance and the energy efficiency operating point scale with the peak power capability of the PA and the used IBO. On the other hand, the practical impact of the EVM on the user data rate depends both on the UE's SINR condition, as well as on the modulation order and code rate utilized for transmission.

While there are distortion-free PAPR reduction solutions that could be used in Massive MIMO base stations, as explained in Section 3.3, some level of distortion can be expected to be tolerable by UEs with no or negligible impact on the achievable data rates. The influence of distortion is explored first considering isolated single cell, based on simulation results covering the BLER performance in Section 5.3.1 and the impact of unequal power allocation in Section 5.3.2. The extension of the analysis to multi-cell network, covered in Section 5.3.3, includes consideration of the balance between the achievable data rates and the energy consumption of the base station.

If usage of distortion-based PAPR reduction is considered to be beneficial from the implementation perspective (e.g., due to relatively low complexity) identification of the combinations of spatially multiplexed UEs that should be avoided could be done by using the EVM performance prediction as an input for distortion-aware scheduling, as demonstrated in Section 5.3.4.

The basis for this part of the research is the original work of the PhD candidate published in [62] (co-authored with the supervisor and the assistant supervisor). While the vast majority of the simulation results presented in Section 5.3 were not published before, some of the results from [62] were reused, as explicitly indicated in the text.

5.1. Preliminaries

Apart from the SE limitation resulting from the channel capacity bound, there are additional practical constraints on the achievable DL data rates in 5G NR transmission, as outlined in

Section 5.1.1. Specifically, the additional SE limitation originating from the maximum configurable modulation order and code rate, as well as the peak user data rate limit resulting from the number of resources available for DL data transmission. When considering the cell throughput, the apparent increase in the SINR due to array gain is offset to some extent by the need to distribute the TX power between users.

The introduction of the wideband SNR concept in Section 5.1.2, which is highly relevant when considering the BLER performance in conventional MIMO transmission over NLoS channels, is followed by a description in Section 5.1.3 of the channel hardening phenomenon which leads to almost deterministic behavior of fading channels.

5.1.1. Practical Limitations of the Achievable Data Rates

The maximum achievable SE, bounded by the channel capacity as explained in Section 2.1, depends on the received powers of the wanted signal S , interference I , noise N , and distortion D as follows:

$$SE = \log_2 \left(1 + \frac{S}{I + N + D} \right) = \log_2(1 + \text{SINDR}) \quad (5.1)$$

In practice, the achievable SE does not increase indefinitely when the SINDR conditions improve. The increase of the SE is additionally limited by the maximum modulation order and code rate supported for data transmission. For Release 18 of 5G NR [70, Sec. 5.1.3], the ceiling on the DL data rate per RE of about 9.3 bit/s is achievable with 1024QAM at 948/1024 code rate, for each of the MIMO transmission layers. Moreover, the limitation of the SE may be additionally influenced by the capabilities of the UE to support a given modulation order. For example, with 256QAM the limitation on the achievable data rate reduces to about 7.4 bit/s per RE.

The relation of the SE limit resulting from a per RE capacity of a point-to-point deterministic AWGN channel, following (5.1), is visualized in Figure 5.1 together with the limits on the SE resulting from 1024QAM and 256QAM using 948/1024 code rate.

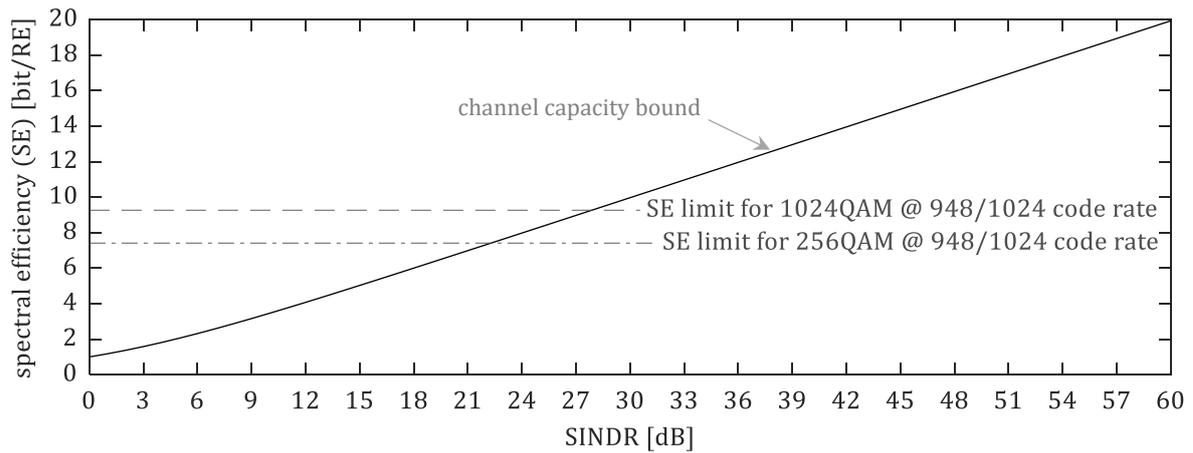


Figure 5.1: The increase in the spectral efficiency (SE) per resource element (RE) with increasing signal-to-interference-plus-noise-plus-distortion ratio (SINDR), bounded by the channel capacity, is limited in practice by the maximum supported modulation order and code rate.

The SE limitation resulting from the maximum supported modulation order and code rate, exemplified in Figure 5.1, suggests that increasing the SINDR above a certain level is not expected to result in increased data rate per RE, in a deterministic AWGN channel. However, that limiting SINDR level is in practice somewhat higher than the crossover point of the SE limitation resulting from the channel capacity bound and SE limits introduced by modulation order and code rate visualized in Figure 5.1 (i.e., about 28 dB for 1024QAM and about 22 dB for 256QAM), since usage of symbols from a discrete QAM constellation leads to a performance gap compared with the Gaussian input assumed in the Shannon's channel capacity limit. Unless stated explicitly otherwise, all the data rates calculated in this chapter utilize the assumption of a Gaussian input, which leads to optimistic relation of the SE with the SINDR.

The peak data rate per UE, for each of the MIMO layers, is limited by the number of REs that can be allocated for the DL transmission of data signals. In 5G NR, the total number of REs available for transmission depends on the base station channel bandwidth (divided into per SC sub-channels) and the configured spacing between SCs [65, Sec. 5.3.2]. For example, a 20 MHz 5G NR channel (NR20) can utilize up to 1272 SCs (when 15 kHz SC spacing is used), while a 100 MHz 5G NR channel (NR100) has up to 3276 SCs (with 30 kHz spacing between SCs). With 14 OFDM symbol per slot, this leads to a total of 17 808 REs per 1 ms for NR20 and 45 864 REs per 0.5 ms for NR100. Since not all the REs are available for DL data transmission (due to signaling overhead, as well as the UL time allocation in case of TDD operation), the achievable data rate per UE is a portion of the

theoretical maximum rate (calculated assuming all REs are available for data transmission), resulting in a ceiling of about 165 Mbit/s for the NR20 and about 849 Mbit/s for NR100.

With Massive MIMO base station, the limit on the cell throughput (i.e., the sum data rate of all the UEs in the cell), resulting from the SE limitation based on the maximum usable modulation order and code rate, scales linearly with the number of spatially multiplexed users, as demonstrated in Figure 5.2. In practice, it is unlikely that all the UEs will be in SINR conditions resulting in reaching the SE limit. In this case, the achievable cell throughput increases with increasing the number of UEs K (increasing the limit on the achievable sum data rate) and the number of antennas M (improving the SINRs by M/K in the best case, since while the array gain increases by M , the TX power is shared by K users).

The relation between the maximum cell throughput for NR100 and the number of UEs is presented in Figure 5.2 with the dependence on the dimensioning of the antenna array, assuming 80% of REs allocated for DL data transmission to UEs at 0 dB SINR (i.e., when $M = K$) and negligible level of distortion. In practice, the achievable sum data rate will depend on the UE SINR conditions, the power allocation strategy, and the performance of the IUI mitigation attained by the utilized precoding approach [9, Sec. 3.5].

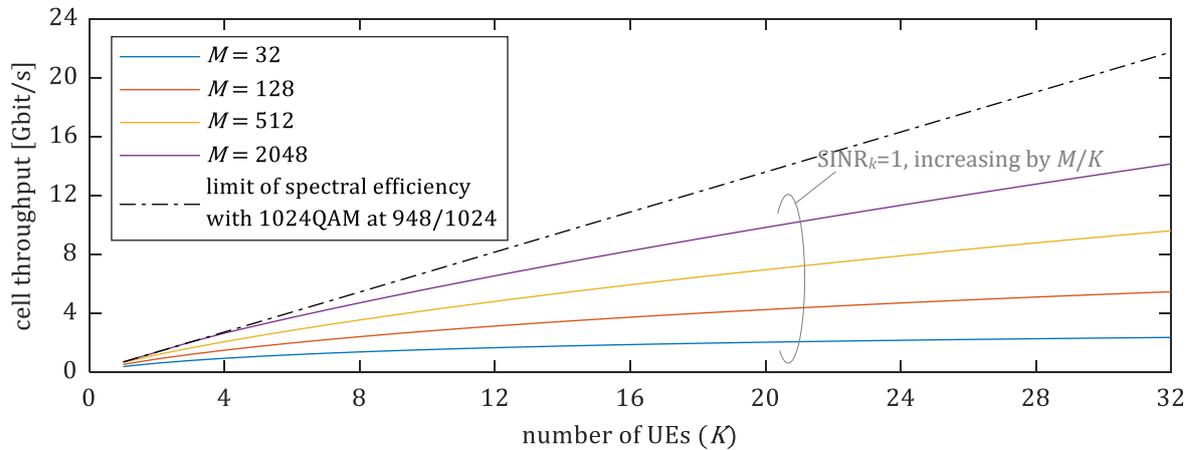


Figure 5.2: Maximum cell throughput for 100 MHz 5G NR carrier for a range of antennas M in relation to the number of UEs K , each with SINR of 1 when $M = K$. Assuming 80% resources available for downlink data transmission and spectral efficiency limited by 1024QAM at 948/1024 code rate.

5.1.2. Wideband Signal-to-Noise Ratio

When the inter-cell interferences can be neglected and the intra-cell interferences between UEs are sufficiently reduced by MU-MIMO precoding, the conditions that the UE is in can be described by the relation between the expected received signal power and the RX noise.

Considering a single UE, the average signal-to-noise ratio (SNR) for the j -th SC can be described, following [1, Eq. 1.7], as

$$\text{SNR}_j = \frac{p_j \mathbb{E} \{ \|\mathbf{h}_j\|^2 \}}{\sigma_j^2} = \frac{p_j \beta_j}{\sigma_j^2} \quad (5.2)$$

where p_j is the TX power allocated to the j -th SC, $\mathbb{E} \{ \|\mathbf{h}_j\|^2 \} = \beta_j$ is the average channel gain for the j -th SC, while σ_j^2 is the RX noise power.

Since the wideband channel covered by the OFDM carrier can be generally considered to be frequency-selective, the received power level is expected to vary over SCs (which effectively divide the wideband channel into multiple sub-channels). This leads to SNR variation over frequency, as exemplified in Figure 5.3 for single realization of multi-path LoS and NLoS channels in SISO transmission using NR20 carrier, when equal TX power is allocated to all SCs.

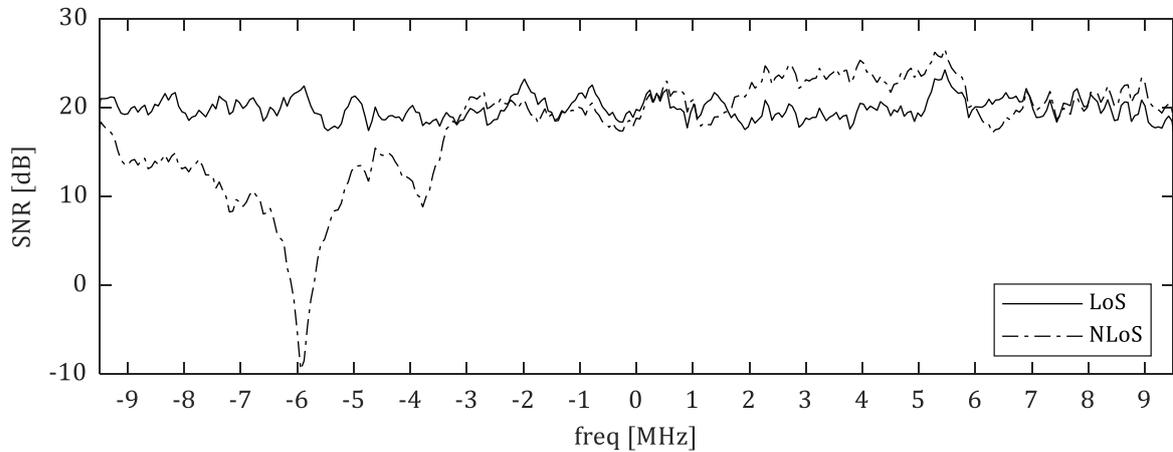


Figure 5.3: Examples of frequency distributions of SNR for an equal level of 20 dB wideband SNR.

A single wideband SNR can be calculated per UE, as the mean of the SNRs over all SCs. The examples for LoS and NLoS channels presented in Figure 5.3 have the same wideband SNR of 20 dB, with significant variations over frequency especially visible for the NLoS channel. Clearly, using multiple SCs for transmission of a single block of data in a coded transmission, with the same modulation order and code rate applied to each used SC, may require a different criterion than a wideband SNR for selection of the appropriate modulation and coding scheme (MCS). Hence, in the link-to-system mapping there is a concept of an effective SNR, with different methods used to calculate this averaged SNR based on the SNRs available per SC.

However, for a Massive MIMO transmission, the effective SNR may not necessarily be that far from the wideband SNR, as a result of the phenomenon of channel hardening, covered in the following section.

5.1.3. Channel Hardening

While in SISO transmission the channel gain can noticeably vary between fading realizations and over SCs, as demonstrated in Figure 5.3, in MISO transmission from Massive MIMO base station the phenomenon of channel hardening leads to a reduction of the variance of the channel gain [1, Sec. 2.5.1]. As a result, the gain of a given realization of a fading channel is getting close to its mean value when the number of antennas $M \rightarrow \infty$. This means that the probability of deep fades is reduced, since such fades would need to occur in the same SCs simultaneously in the channel responses from many antennas, hence the selection of frequency resources for scheduling of a given UE becomes less relevant.

This kind of behavior, where deep fades that occur in some of the antennas do not cause a significant degradation from the MISO channel perspective, is exemplified in Figure 5.4. For one of the three antennas, randomly selected out of 64 antennas used for transmission in a spatially correlated NLoS channel, a deep fade is clearly visible in the presented fading realization. However, the mean channel gain $\|\mathbf{h}_k\|^2$, which provides averaging over all antennas per SC, is relatively flat compared with the channel responses of single antennas. This means that a UE could be assigned any of the available SCs, since the narrowband SNRs would not significantly deviate from its wideband SNR.

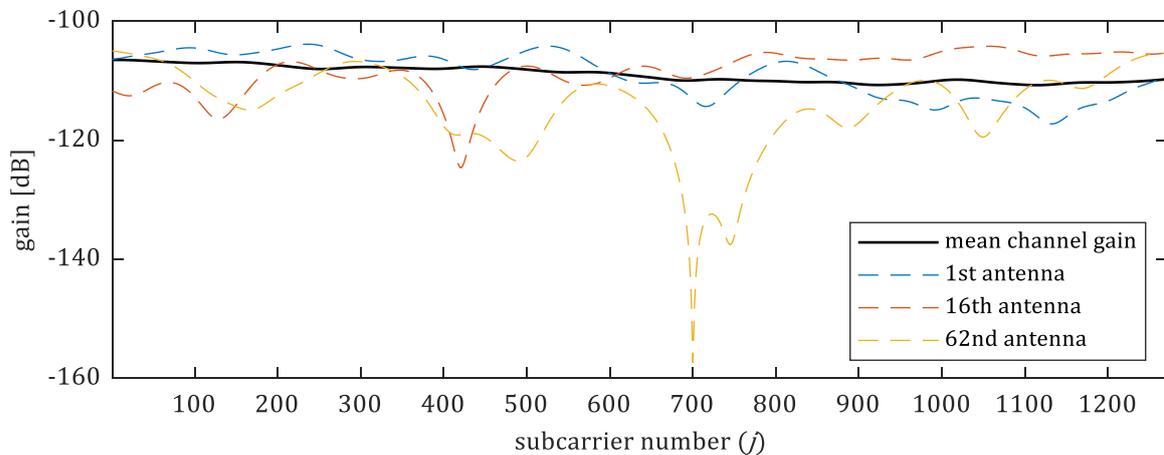


Figure 5.4: Example of the mean channel gain and gains of the channel responses of 3 antennas out of 64.

A channel hardening measure, based on [71, Eq. 16], can be defined for each of the J SCs as follows:

$$\frac{\text{Var}\{\|\mathbf{h}_k\|^2\}}{(\mathbb{E}\{\|\mathbf{h}_k\|^2\})^2} \quad (5.3)$$

where more hardening leads to (5.3) approaching 0.

An example of comparison of the level of channel hardening in spatially correlated NLoS channel, depending on the number of antennas M is visualized in Figure 5.5, based on

100 fading realizations. It is clearly visible that an increase in the number of antennas leads to more channel hardening (i.e., the measure from (5.3) approaches 0). The differences observable between the results for 64 and 128 antennas are relatively small, which shows that the convergence of the channel hardening measure towards 0 slows down significantly. In this case, even for 128 antennas, only a partial channel hardening can be observed, as expected for spatially correlated channels [1, Sec. 2.6.1].

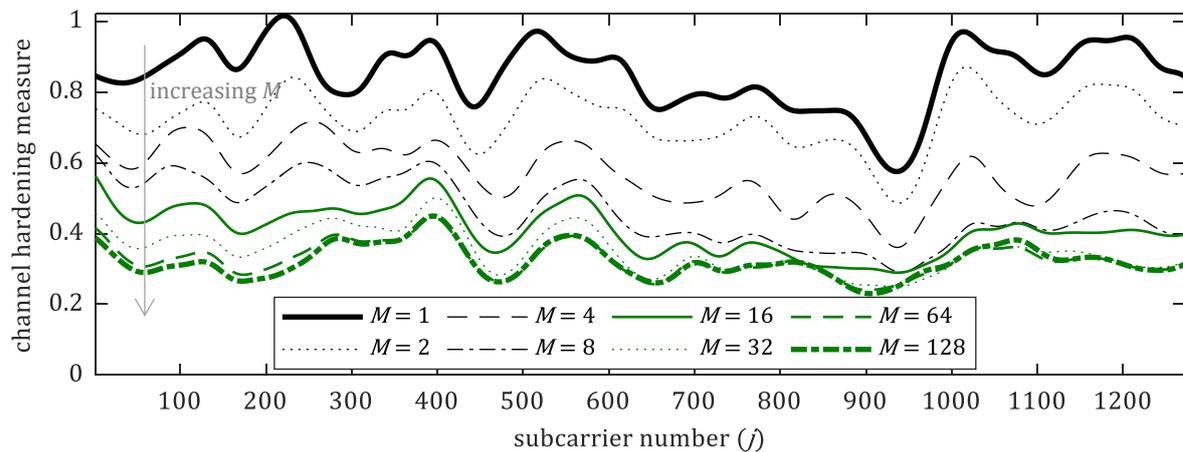


Figure 5.5: Channel hardening measure, from (5.3), per subcarrier for simulated spatially correlated NLoS channels improves for higher number of antennas M . Based on a figure originally published in [62] (co-authored by the PhD candidate).

5.2. Methods

There are two different approaches followed here to reflect the impact of distortion on the achievable data rates. The first method is to characterize the BLER performance for a given modulation order and code rate, based on the LLS results. This approach is followed in the analysis described in Section 5.3.1, focused on capturing the impact of channel hardening, by considering separately the performance with Gaussian noise added to either TX or RX side of the link. The second method, using the SDR calculated from the EVM performance (either estimated or based on the results of simulation) as an input for the SE calculation in (5.1), which is used as a basis for the user data rate and the cell throughput calculations according to the description in Section 5.1.1. This method is used in all the other studies covered in Section 5.3.

In order to calculate the SE, apart from the SDR also the SINR is required. Here, two different strategies are followed when it comes to the SINR performance. When focusing specifically on particular aspect of distortion, like the behavior with unequal power allocation in Section 5.3.2 or the relevance of the radiation pattern of nonlinear distortion in

Section 5.3.4, simplified assumptions on the relative differences of SINRs between UEs are made and the data rate performance is characterized over a wide range of SINRs conditions. The second approach used to characterize the impact of distortion in multi-cell network in Section 5.3.3 is based on numerical simulations, as described in more detail in Section 5.2.1.

5.2.1. Reflecting the Impact of Nonlinear Distortion in a Multi-Cell Network

The performance in multi-cell network is captured here using the same approach as outlined in [1, Sec. 4.1.4], based on Monte Carlo numerical simulations.

The simulation code shared together with [1] to generate the data for characterization of the average DL sum SE for different precoding schemes, which results are visualized in [1, Fig. 4.18], is utilized here as a basis. The impact of distortion is reflected by modifying the SE calculation in [1, Eq. 4.38] to base on the SINDR instead of the SINR. The SINDR for the k -th UE is calculated as

$$\text{SINDR}_k = (\text{SINR}_k^{-1} + \text{SDR}^{-1})^{-1} \quad (5.4)$$

with both the SINR_k and the SDR in a linear scale. The SINR_k is the result of the numerical simulation, one for each of the 500 simulated channel realizations for a given drop of UEs. The SDR is calculated based on the best-case EVM performance prediction model in (4.19), which depends on the per antenna EVM, the number of UEs, and the number of antennas.

The example of a 16-cell network, deployed in a square pattern with a wrap-around, defined in detail in [1, Sec. 4.1.3] is followed here, with the pilot reuse factor of 1. The same parametrization of the example as in [1, Fig. 4.18] is used (10 UEs dropped randomly per cell, with equal power allocation of 20 dBm per UE; the Gaussian local scattering model, defined in [1, Eq. 2.23] is employed).

5.3. Simulation Results and Discussion

First, the impact of TX distortion and RX noise on the BLER performance, depending on the number of antennas, is compared in LLS for single UE transmission. It is followed by consideration of the impact of unequal power allocation on the achievable DL data rates. Next, the performance in multi-cell network, using both single- and multi-cell precoding solutions, is explored in Section 5.3.3. Finally, a demonstration of the performance improvement possibility enabled by utilizing the precoding-based EVM prediction model as an input for making distortion-aware scheduling decisions is provided in Section 5.3.4.

5.3.1. Block Error Rate Performance with Transmitter and Receiver Noise

The quantization error, modeled as an additive complex Gaussian noise uncorrelated between antennas, can be expected to be radiated with the array element pattern, as explained in Section 4.5. Therefore, the received power of the related distortion is not expected to increase with increasing number of antennas M (i.e., provided that the sum output power is kept constant irrespective of M). Since the wanted signal power benefits from the array gain, an improvement of the transmission quality experienced by the UE is expected, increasing the SNDR by up to M . This allows to increase the UE data rate, provided that it is not limited already by the maximum usable modulation order and code rate, as explained in Section 5.1.1.

To compare the impact of the RX noise and the impact that the part of residual TX impairments which could be represented relatively accurately by the uncorrelated Gaussian noise (like the quantization distortion potentially) has on the achievable user data rate, a simulation campaign using varying levels of either noise power N at the RX or the per antenna distortion power at the TX side D_m (resulting in a variation of the received power of distortion D) for a MISO transmission over the CDL-C channel (i.e., spatially correlated frequency-selective NLoS) was executed in LLS. For each number of antennas M , the maximum power level of the added Gaussian noise (i.e., the minimum wideband SNDR—either SNR or SDR in this case) allowing to reach BLER not higher than 10% (during a transmission of 100 slots) for 256QAM at 948/1024 code rate, was identified.

It can be observed in Figure 5.6 that an increase in M leads to a relaxation of the ratio between the signal power (S for the received and S_m for the per antenna) and the power of the added Gaussian noise (representing either the noise power N at the RX or the per antenna distortion power D_m at the TX) required for successful transmission (i.e., allowing to tolerate higher powers of either noise or distortion).

In the verified scenarios, the power ratio required for the case of RX noise was considerably relaxed, going down from about 33.9 dB for $M = 1$ to about 27.1 dB for $M = 64$, reaching the SNR level required for 10% BLER for $M = 1$ in CDL-D (LoS) channel (which was however still about 1.4 dB above the SNR that was required to reach 10% BLER in a free-space LoS AWGN channel), as demonstrated in [62]. This suggests that thanks to the increased level of channel hardening, leading to reduced probability of deep fades as explained in Section 5.1.3, with increasing number of used antennas M we could expect a similar BLER performance for NLoS and LoS channels. It is worth to point out that the main

part of the observed reduction in the needed SNR (about 6.0 dB out of a total of about 6.8 dB) was achieved when increasing the number of antennas from 1 to 8, with diminishing gains observed with each step up (about 0.1 dB when going from 32 to 64 antennas). This suggests that already with the number of antennas used in the initial Massive MIMO deployments [11], a major part of the potential benefits coming from the channel hardening could be obtained, which is in line with the observations based on the analysis of the channel hardening measure in Figure 5.5.

When considering the results for the TX distortion, the tolerable increase of the per antenna power of distortion D_m with increasing the number of antennas M converged to about 3 dB for each doubling of antennas. This is in line with the expected difference in the received power of the fully coherently combined wanted signal (i.e., $S = M^2 S_m$) and the incoherently combined distortion (i.e., $D = M D_m$). When compensating for the expected difference of array gain in the received power, an improvement in the BLER performance, attributed to the channel hardening (similarly to the case of the RX noise) can be observed in Figure 5.6 for the TX distortion as well. In that case about 3.9 dB reduction in the required SDR was achieved, with about 3.5 dB relaxation achieved already with 8 antennas, down to the level of about 26.3 dB which was less than 0.1 dB away from the SDR that was required for 10% BLER for $M = 1$ in CDL-D (LoS) channel, as demonstrated in [62].

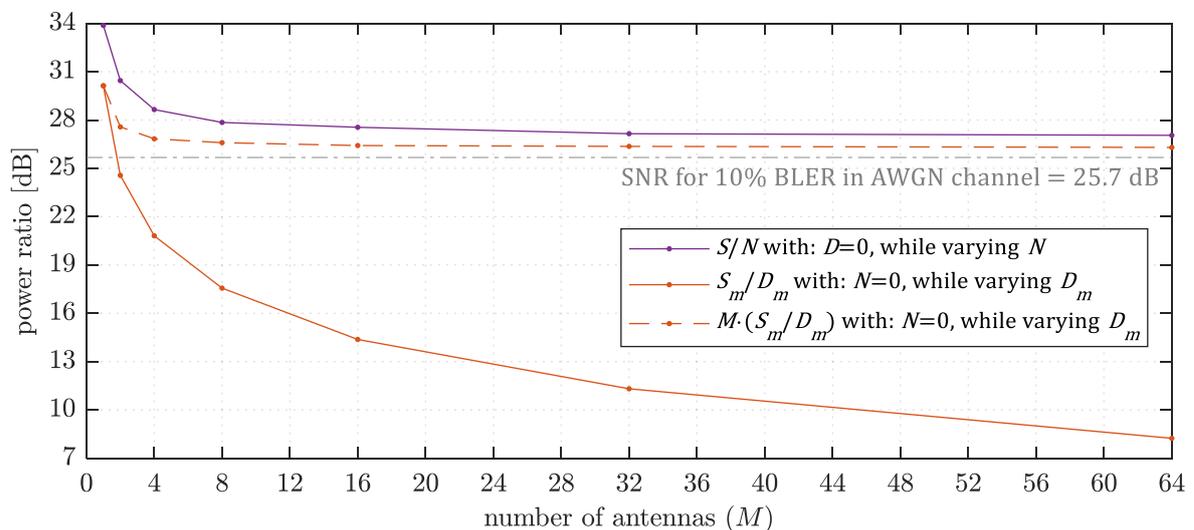


Figure 5.6: Minimum ratios between the wanted signal power (S for received, S_m for per antenna) and the power of the additive Gaussian noise (representing either the noise power at the receiver N or the per TX antenna distortion power D_m) required for 10% BLER in CDL-C (NLoS) channel depend on the number of antennas M . Based on a figure originally published in [62] (co-authored by the PhD candidate).

The minimum SNR of 25.7 dB required in the AWGN channel for 10% BLER with 256QAM at 948/1024 code rate, visualized in Figure 5.6, was obtained as a result of

simulation over 100 slots. While it is not excluded that a more accurate result could be achieved with an order of magnitude longer simulation time, this is roughly a level that could be expected based on the performance evaluation of 256QAM in Vienna SLS in [72] as well as in a number of 3GPP studies for 5G NR. For example, in the results shared in [68], for a code rate of 948/1024 the SNR required for 10% BLER for 256QAM seems to about 25.3 dB, while for 1024QAM about 30.9 dB, with 1% BLER level reached with less than additional 0.5 dB increase in the SNR. Note that to achieve comparable levels of channel capacity, about 3 dB lower SNR is needed when a Gaussian input is assumed.

Note that in this verification, the effect of the SNR improvement with increasing the number of antennas was not explicitly visible since the focus was on the relation of the SNR and the BLER performance. Notice what the shared results actually mean from the perspective of a UE experiencing a channel quality described by a given SNR level. For example, roughly the same SNR was needed for the same BLER for 32 and for 64 antennas. This means that the UE could have about 3 dB higher power of the RX noise when the number of antennas is doubled, since the received signal power will increase by 3 dB due to increase in the array gain.

5.3.2. Data Rate Impact of Unequal Power Allocation

A comparison between the impact of equal and unequal power allocation on the EVM performance was explored in Section 4.7.5, for residual TX impairments modeled either as nonlinear distortion or uncorrelated noise. Using the same example of 4 UEs (located at 50, 75, 100, and 500 m distances from the base station), the relation between the EVMs and achievable data rates is considered here.

When the strategy of equal power allocation is used, the UEs with different path losses experience different SNRs. Assuming SNRs aligned with the differences in the free-space path loss in the example from Section 4.7.5, the UEs further away from the base station have SNRs lower by about 3.5 dB, 6.0 dB, and 20.0 dB in relation to the UE that is the closest. On the other hand, when the strategy of power allocation with the target to achieve comparable QoS is used, all the UEs experience the same SNRs as a result. Clearly, compared to equal power allocation, this leads to a reduction in SNR for some of the UEs and increase in SNR for others, with the intention to enable provision of comparable data rates. In the considered example, that equal SNR was about 14.3 dB lower than the SNR achieved by the UE closest to the base station when equal power distribution is used.

Since unequal power distribution over UEs may lead to EVM performance degradation, as demonstrated in Section 4.7.5, the SNDRs resulting from the combination of noise and distortion powers should be considered when evaluating the impact of the power allocation strategy on the achievable data rates. The achievable sum data rates per RE, depending on the SNR (in case of equal power allocation) of the UE closest to the base station, presented in Figure 5.7, consider the distortion power estimates resulting from the EVM performance of ICF with 6 dB CR, shared in Table 4.8. That EVM performance led to about 28 dB SDR for all the UEs in case of the equal power allocation, for both 32 and 128 antennas. While for the equal SNR strategy, for $M = 32$ the resulting SDRs were in the range from about 18 dB to about 25 dB and for $M = 128$ in a somewhat improved range from about 21 dB to about 26 dB.

When the limitation related with the maximum configurable modulation order and code rate is considered, the achievable data rate saturates at the related level when the effects of distortion are ignored, as visible in Figure 5.7. However, when the impact of nonlinear distortion is considered, the sum data rate may reach a lower limit with increasing SNR, imposed by the SDR. This is especially visible for the equal SNR strategy, since the EVMs obtained with ICF for the selected CR leads to SDRs that are lower in this case than the about 28 dB required for 1024QAM at 948/1024 code rate, as explained in Section 5.1.1.

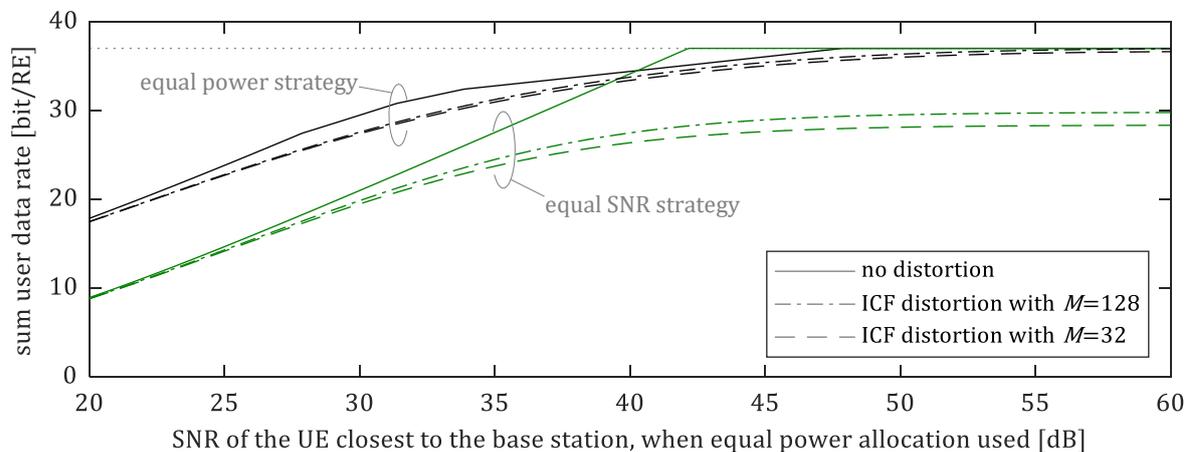


Figure 5.7: Comparison of the sum user data rates for 4 UEs located at different distances from the base station (50, 75, 100, and 500 m) for 2 different power allocation strategies, with no distortion and with iterative clipping and filtering (ICF) distortion from Table 4.8, for linear array with different number of M antennas.

Since the purpose of the equal SNR strategy is not to increase the achievable sum data rate, but it is instead to provide fairness by allocating power in a way to achieve comparable QoS, a detailed look at the per UE achievable data rates is required. Clearly, that goal may be achieved when the impact of distortion is ignored, as visible in Figure 5.8(b). However,

the consequence of the unequal EVM performance over UEs, observed in case of the equal SNR strategy for both uncorrelated Gaussian noise and nonlinear distortion cases in the results shared in Section 4.7.5, is that an equal QoS may not be easily achievable when distortion is considered. Moreover, it is not excluded that the achievable data rate of users that were originally in better SNR conditions will saturate at lower levels than for the UEs that experience worse channel quality, as visible when comparing the results in Figure 5.8(a) with Figure 5.8(b).

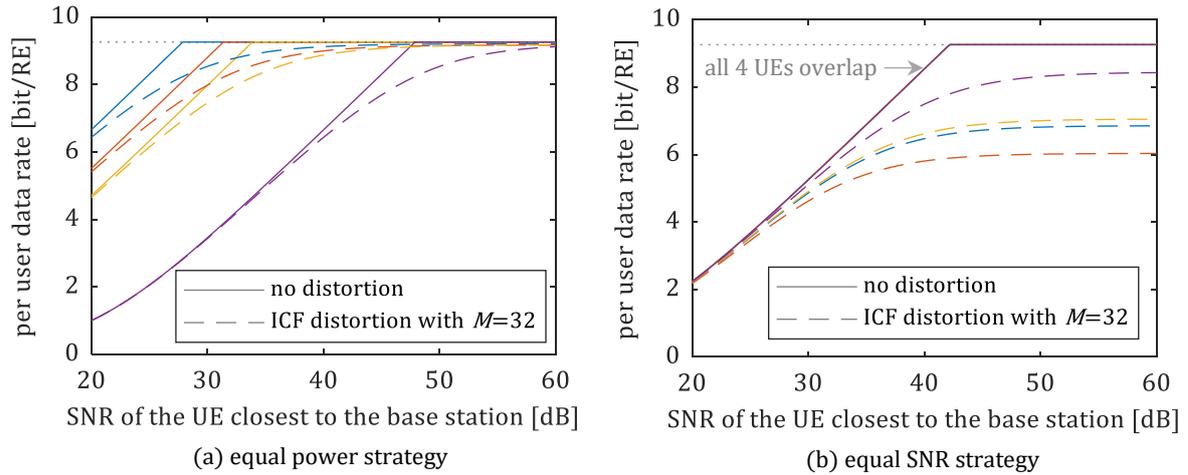


Figure 5.8: Per user data rates (each color represents specific UE) for two different power allocation strategies, with no distortion and with iterative clipping and filtering (ICF) distortion, as a function of the SNR of the UE closest to the base station when equal power allocation is used.

5.3.3. Balance Between the Achievable Data Rates and Energy Consumption in Multi-Cell Networks

The analysis of BLER performance, covered in Section 5.3.1, as well as the consideration of the impact of the power allocation strategy, described in 5.3.2, relied on using simplified assumptions on the UE channel quality, in order to clearly demonstrate the specific aspects explored by these dedicated studies. For the same reason, the BLER and data rate performance impacts were considered ignoring the effects of interference (both intra- and inter-cell). Reliance on the SNR as a measure of the channel quality perceived by UEs, may represent well transmission in an isolated cell, provided that the intra-cell interference is handled by precoding. In order to capture the complexity related with transmission in a cellular network, the next step is to extend the analysis to cover multi-cell scenarios and deal with the potentially performance-limiting impact of the inter-cell interference.

The system-level aspects of DL transmission in cellular networks are handled here in a way aligned with the approach taken in [1], as explained in more detail in Section 5.2.1.

Assuming there are $K = 10$ UEs randomly dropped in each of the 16 cells of the network, the best-case EVM performance prediction model can be used to provide limits on the SDRs, and by extension, on the per UE achievable data rates. The actual impact on the data rates will depend on the SINR conditions the UEs were originally in. To understand what statistics of the SINR could be expected for the considered scenario, depending on the number of antennas M in each of the base stations, numerical simulations were performed for $M = 32$, $M = 128$, and $M = 512$. The CDFs of SINRs, presented in Figure 5.9, clearly show improvement of the SINR for all transmissions with increasing the number of antennas. Additionally, the simulation results seem to confirm that usage of multicell minimum mean-squared error (M-MMSE) precoding can be expected to outperform the precoding solutions that do not handle inter-cell interference, like the regularized zero forcing (RZF) visible in Figure 5.9.

Note that the SINR statistics in Figure 5.9 are based on 500 channel realizations for a single random drop of UEs. While different verified drops led to somewhat different results, the observations related with the dependence on the array size and relation between precoding approaches still held.

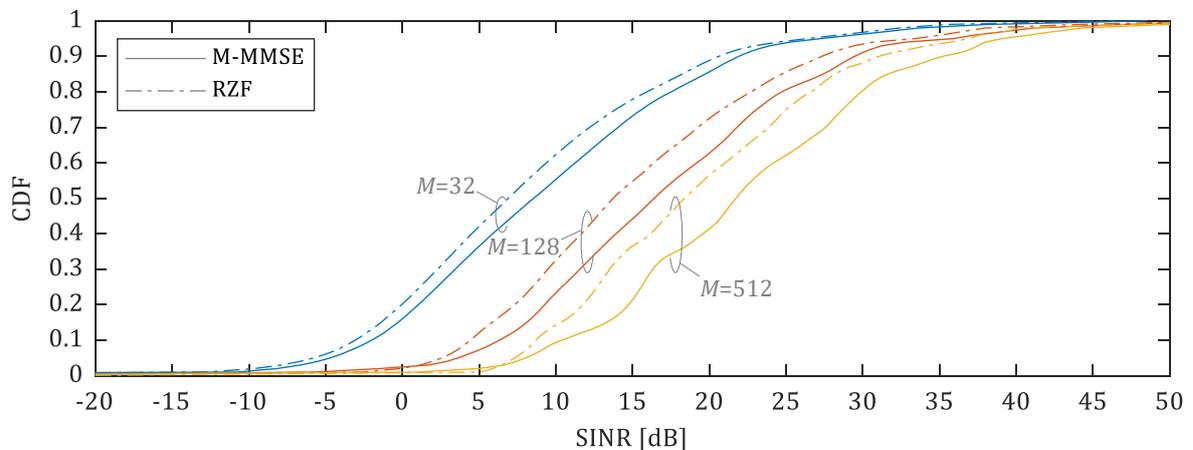


Figure 5.9: Cumulative distribution functions (CDFs) of signal-to-interference-plus-noise ratio (SINR), with multicell minimum mean-squared error (M-MMSE) and regularized zero forcing (RZF) precoding, for 16-cell network with 10 UEs and base station with M antennas in each cell.

Clearly, different levels of distortion can be expected to influence the cell throughput differently. When comparing the CDFs of SINR from Figure 5.9 with the expected levels of EVM from the best-case EVM performance prediction model in (4.19) it becomes clear that both dependence on the number of antennas (leading to limits related with uniformly radiating distortion) and the per antenna EVM itself (depending on the peak power limit in case of distortion-based PAPR reduction or saturation in the PA) is relevant.

First, the relation between the sum data rates per RE, averaged over cells, is considered for M-MMSE precoding in relation to the number of antennas. This replicates the results of [1, Fig. 4.18] as a reference with excluded effect of distortion. The results with distortion were obtained by considering SDRs calculated based on the best-case predictions model in (4.19), for different levels of per antenna EVM. The summary of the results from 10 drops of UEs is provided in Figure 5.10. For this comparison the limitation on the achievable SE resulting from the maximum configurable modulation order and code rate was omitted, to provide a clear comparison point with the results shared in [1, Fig. 4.18]. Note that the same seed was used for configuration of the random number generator for each of the EVM performance levels, to obtain repeatable results allowing to focus on the differences coming solely from the reflected impact of distortion.

With $\text{EVM}_{\text{ant}} < -30$ dB, the resulting data rates were virtually overlapping with the results without distortion, hence they were not visualized in Figure 5.10. When the per antenna EVM was reducing from -30 dB down to 0 dB the degradation of the average sum data rate per cell was getting larger. This could be expected especially for EVM_{ant} of -10 dB and 0 dB, when judging what portion of UEs could have SINR high enough (taking the SINR distributions in Figure 5.9 as a reference) to be impacted by the low resulting SDR, which was improved in the best-case by about 7.2 dB compared to the per antenna EVM.

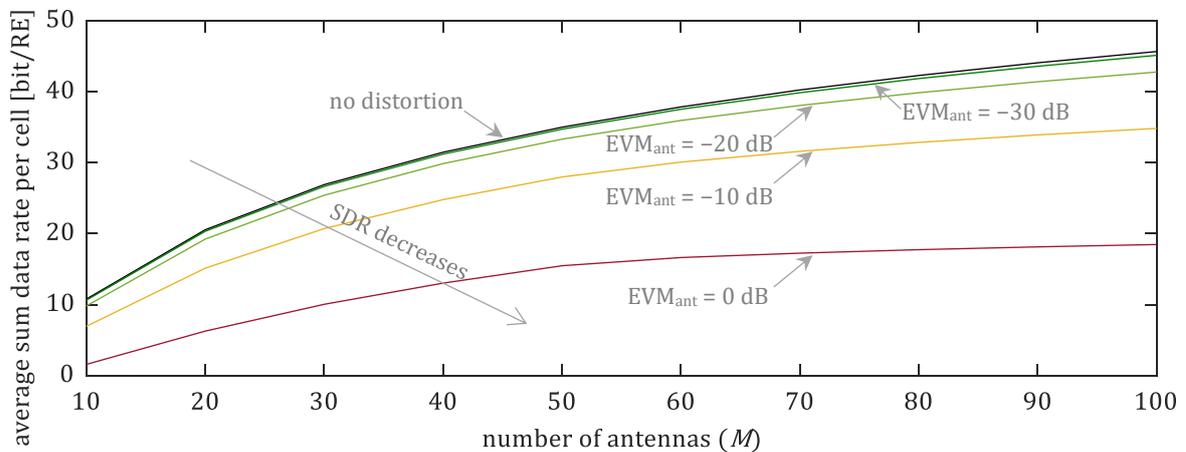


Figure 5.10: Sum data rate per resource element (RE) for M-MMSE precoding, depending on the number of antennas, averaged over cells for a range of per antenna EVM levels.

For example, with $M = 100$ antennas, the results show a minor drop of only about 1% in the average sum data rate for EVM_{ant} of -30 dB compared to the results without distortion. However, when the per antenna EVM was increased, a visible degradation of the achievable data rates was observed (i.e., about 6% for -20 dB, 24% for -10 dB, and 60% for 0 dB). This proves that the effects of the residual TX impairments cannot be generally ignored,

even when the number of antennas goes to infinity, since the portion of the energy of nonlinear distortion radiating towards UEs is determined by the number of UEs that the base station is transmitting to in a given slot, as clearly reflected in (4.19).

The results shared in Figure 5.10 suggest that relaxation of the requirements on the per antenna EVM (which relates with the signal quality requirements that the components in the TX processing chain need to fulfill and the energy efficiency operating range for these components) leads to increasing performance gap to the data rates achievable with no distortion when the number of antennas on the base station side is increasing. This is in line with the expectations based on the SINR cell geometry shared in Figure 5.9, which reveals a shift of the CDF towards higher range of SINR values with increasing number of antennas, resulting in more UEs impacted by distortion.

Assuming that the nonlinear distortion comes fully from the PAPR reduction using ICF, the relation between the achievable data rates and energy consumption can be established by connecting the per antenna EVM performance with the respective levels of IBO required at the input of PAs. Using the per antenna EVM performance dependence on the peak power limitation by ICF, covered in Section 4.7.1, the relation between the IBO (including correction for the power loss caused by PAPR reduction, to have comparable levels of output power for each of the energy efficiency operating points) and EVM is presented in Table 5.1.

The sum data rates per RE averaged over cells, obtained from the simulation for M-MMSE, RZF, and MR precoding are presented in Table 5.1. The limit on SE of about 9.3 bit/RE, resulting from the highest configurable modulation order and code rate (i.e., 1024QAM at 948/1024) was used in this case, leading to a ceiling on the sum data rate per cell of about 93 bit/RE with 10 UEs. Clearly, with a larger IBO (resulting in lower EVM) generally higher sum data rate could be expected, however the per UE data rates are limited not only by SDRs but (more importantly) by their channel conditions reflected in SINRs, leading to obtaining only a portion of the max achievable data rate even for IBO of 10 dB. In that case, with SDR of nearly 60 dB only about 43% of the peak data rate was obtained. Note that 5% of samples were used in this case for pilots, reducing the sum data rates.

When decreasing the IBO from 10 dB down to 3 dB, a similar ratio of noticeable reduction in the sum data rate was observed for M-MMSE and RZF (i.e., about 19% and about 18%, respectively) however only about 7% reduction was observed for MR precoding. This results from much lower SINR achievable with MR, due to the impact of both intra- and inter-cell interferences.

Table 5.1: The relation between the input back-off (IBO) and the per antenna EVM, the resulting signal-to-distortion ratio (SDR) for 64 antennas M and 10 users K , and the sum data rate per resource element (RE) averaged over cells for M-MMSE, RZF, and MR precoding.

IBO [dB]		3	4	5	6	7	8	9	10
EVM_{ant} [dB]		-10.5	-14.0	-17.8	-21.8	-27.0	-33.3	-41.3	-51.8
SDR [dB] (for $M=64$ and $K=10$)		17.7	21.2	25.0	29.0	34.2	40.5	48.5	59.0
sum data rate per RE averaged over cells [bit/RE]	M-MMSE	31.9	34.9	37.1	38.6	39.3	39.5	39.5	39.6
	RZF	28.6	31.0	32.8	33.9	34.4	34.6	34.7	34.7
	MR	16.6	17.2	17.6	17.7	17.8	17.8	17.8	17.8

Clearly, the quantitative impact of IBO selection depends on the one hand side on the array geometry and the number of scheduled UEs. But on the other hand, it depends on the SINR distribution in the cell, that is influenced by the array geometry, parameters of the deployment, and utilized precoding. The M-MMSE precoding (handling the multi-cell interference) clearly leads to higher achievable data rates, making it also the precoding scheme that was most impacted by TX distortion.

The results in Table 5.1 suggest that IBO could be reduced, in order to improve the energy efficiency operating point, provided that some level of reduction in the achievable data rate is tolerable. For example, the 6 dB IBO operating point could be used with about 2% reduction in the achievable data rate for both M-MMSE and RZF, while less than 1% reduction was observed in this case for MR.

The relation between the average DL cell throughput and the energy consumption for a base station can be established based on the simulation results presented in Table 5.1 and the relation between the IBO and energy efficiency. Here, to provide a quantifiable example, a 50% energy efficiency with 0 dB IBO is arbitrarily assumed, reducing by 2 percentage points for each 1 dB increase in the IBO (same as in [73]). The resulting relation between the power usage, calculated assuming 40 dBm sum output power (which is 10 dB higher than used in the simulated example, to represent a situation with larger inter-site distances than used in the simulation), and the achievable average cell throughput is visualized in Figure 5.11 for M-MMSE and RZF precoding. In the lower part of the IBO range (going from 3 dB to 10 dB, with 1 dB step size in Figure 5.11), increase in average cell throughput at the cost of increase power usage can be clearly observed. However, while increasing IBO (mapping essentially in this case to using PA with a higher peak power capability) reduces EVM in the full visualized IBO range, the reduced power of distortion does not translate to noticeable improvement in the cell throughput performance when going beyond about 7–8 dB IBO.

This clearly shows that the additional energy consumed to reduce the distortion to negligible levels was unnecessarily spent in this case.

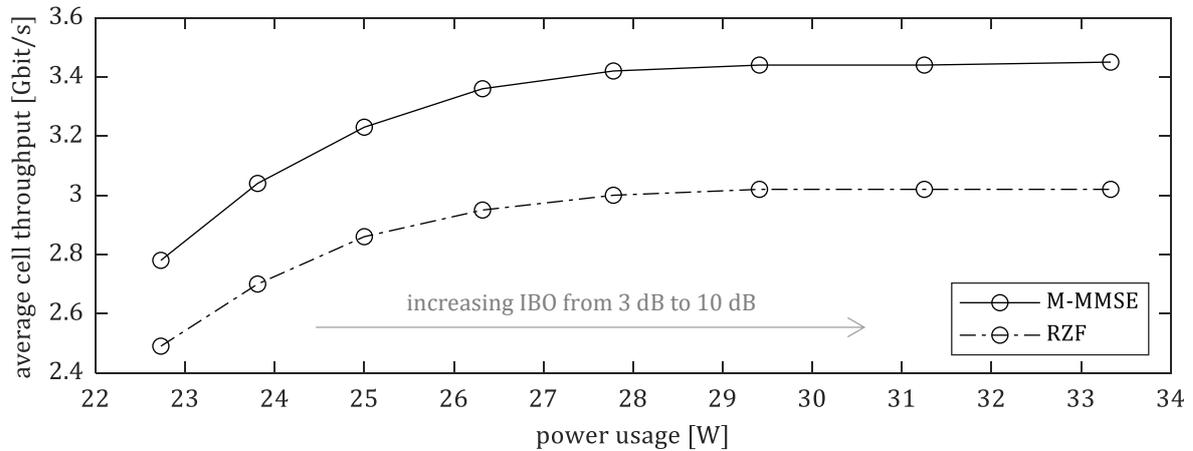


Figure 5.11: Larger input back-off (IBO) leads to higher power usage, but beyond some level it may result in diminishing returns from the perspective of the cell throughput performance.

It was demonstrated in Figure 5.10 that the distortion may negatively affect the achievable DL data rates. The significance of this impact depends on the one hand on the quality of data transmission from the Massive MIMO base station, influenced by the per antenna error level. On the other hand, it depends on the SINR conditions the UEs are in, which was shown in Figure 5.9 to be improving with increasing number of antennas.

While the relation between the SINR and the SE used here leads to optimistic estimates of the achievable data rates (due to assuming Gaussian input in the channel capacity bound, as pointed out in Section 5.3.1), the qualitative characteristic of the dependencies on the impact on data rates when the number of antennas is increased and the per antenna EVM is increased are expected to hold.

The impact of distortion was also optimistic because of using the best-case EVM performance prediction model, which does not take into account the impact of the angular locations of UEs. The potential degradation of the data rate performance caused by spatially multiplexing UEs with unfavorable combinations of angular directions could be avoided by considering precoding-based EVM performance prediction in the distortion-aware scheduling, which is explored in Section 5.3.4.

5.3.4. Performance Improvement Potential of Distortion-Aware Scheduling

It was demonstrated in the previous sections that the residual TX impairments may limit the achievable data rates. This may degrade the user experienced data rates (and consequently

the cell throughput) depending on the relation between the channel conditions of the UEs (represented by the SINRs) and the power of received distortion.

The impact of the part of the per antenna introduced distortion which combine incoherently at the user positions, like the uncorrelated quantization noise covered in Chapter 4, may be reduced by using more antennas on the base station side. However, when energy-efficient Massive MIMO base stations are considered, the per antenna power limitation (resulting from the constraint of the peak power capabilities of the PAs) leads to introducing nonlinear distortion—either by saturation in the PAs or by an explicit PAPR reduction (provided that a distortion-free PAPR reduction is not employed)—which at least partially combine coherently at the UEs even at the limit of infinite antennas.

Since it was established in Section 4.4.2 and confirmed in simulation in Section 4.7 that the EVM performance improvement compared to the per antenna EVM depends not only on the array geometry but also on the relation between the precoding and channel matrices, there is a potential for performance improvement by introducing distortion awareness in user scheduling and link adaptation. If the level of distortion expected to reach the UEs can be estimated with an acceptable accuracy before the transmission, it can be used as an input for making distortion-aware radio resource management (RRM) decisions. For example, this kind of estimate could be used when determining which UEs to spatially multiplex in a given time slot and when selecting the modulation order and code rate applicable to each of the users in that transmission period.

When the antenna array provides a fine enough angular granularity that the potential for receiving a significant portion of energy of the spatially filtered IMPs radiating away from the UEs (see Section 4.4.2.4) is negligible, the simplified best-case EVM performance prediction model (introduced in Section 4.4.2.1) may be used to provide an EVM estimate. Potentially in combination with identification of the portion of spatially filtered IMPs that radiate in the directions given by the precoding weights applied to the wanted data signal, as explored in Section 4.4.2.2. However, when the limited angular granularity leads to unacceptably high inaccuracy of the EVM performance prediction, the precoding-based model proposed in Section 4.4.2.4 could be used instead. In the end, the EVM prediction accuracy, explored in Section 4.7.3 and 4.7.3.3 in relation to the number of UEs and the array geometry for both the proposed models, may influence the performance from the data rate perspective.

When the EVM performance dependency on the spatial locations of UEs may be primarily limited to identification of worst-case combinations of steering angles, described

analytically in Section 4.4.2.3 and explored in simulation in Section 4.7.2, the consideration of distortion awareness in the user scheduling and link adaptation may be relatively straightforward. This could be the case when the dependence of the EVM performance on the location of spatially multiplexed UEs is reflected in relatively steep CCDFs of EVM, which means that majority of potential UE co-scheduling cases lead to a comparable EVM performance. This kind of characteristic was demonstrated in the statistics visualized in Section 4.7.3 and 4.7.3.3, based on both results of the LLS and numerical simulations (using Monte Carlo trials to generate channel matrices and, subsequently, estimate the EVM based on the interaction of precoding and channel matrices).

For example, looking at the statistics of the EVM performance prediction based on numerical simulations presented in Figure 4.21 (covering one million of Monte Carlo trials of sets of UE locations), it is clearly visible that most of the cases are relatively close to the median, while the apparently asymmetric heavy-tailed distribution of EVM have especially problematic outliers with worse EVM performance. This is captured in the comparison of the ranges of EVMs from 10th to 90th percentile (i.e., covering 80% of cases) and from 0th (i.e., the worst-case) to 50th percentile in Table D.3. When the angular resolution of the array was improved, by increasing the number of antennas M from 32 to 128, a narrower range of EVM from 10th to 90th percentile and a wider range from 0th to 50th percentile EVM (except for $K = 2$) was observed. This suggests that the benefit of identifying combinations of UE angular locations leading to relatively poor performance, compared to the typical case, increases with increasing the number of antennas.

Table 5.2: Ranges between percentiles of the EVM performance statistics over sets of UE locations, based on the free-space LoS prediction results from Figure 4.21.

Number of UEs (K)	range from 10th to 90th percentile EVM [dB]		range from 0th to 50th percentile EVM [dB]	
	$M = 32$	$M = 128$	$M = 32$	$M = 128$
1	0.0	0.0	0.0	0.0
2	0.1	0.0	1.2	1.2
3	0.6	0.1	2.5	2.6
4	0.9	0.4	3.2	3.6
6	0.8	0.6	4.0	4.9
8	0.7	0.6	4.0	5.7

Clearly identification of the occasional bad performance cases (e.g., with EVM performance more than 1 dB away from the typical level) may be beneficial. Since it could allow to avoid spatially multiplexing UEs which combination of precoding weights is

expected to result in such a bad performance outlier case and scheduling transmission to these UEs in different time slots instead. Alternatively, it could allow to preemptively reduce the modulation orders and code rates, instead of increasing the block error transmission failure rate.

To demonstrate the potential impact of an occasional relatively poor choice of users to be spatially multiplexed by a distortion-unaware scheduler, the precoding-based EVM prediction model from Section 4.4.2.4 was used to identify a set of UE locations that may lead to radiating a noticeable portion of the spatially filtered IMPs to the directions of users.

A Massive MIMO base station with a rectangular array of $M = 128$ antennas was used. The array was constructed using isotropic sources located in $M_z = 8$ rows and $M_y = 16$ columns, with $d_z = 1.0\lambda$ spacing between rows and $d_y = 0.5\lambda$ spacing between columns, leading to a comparable angular resolution in azimuth and elevation. The steering range was limited to $\pm 50^\circ$ in azimuth and $\pm 25^\circ$ in elevation, while the steering angles were constrained to integer degrees only.

First, angular directions of 2 UEs were randomly selected. The expected radiation patterns of spatially filtered IMPs are visualized in Figure 5.12, down to a -10 dB level from the maximum observed value, for transmission over free-space LoS channel to UEs positioned in the directions marked with crosses. It is visible in Figure 5.12(a) that for 1 UE the peak of the main lobe of distortion pointed to the UE location, as expected. While, when the 2nd UE was added, there were 2 distinct lobes of distortion that can be observed in the range of angles visible in Figure 5.12(b). This led to about 1.7 dB improvement in SDR compared to the 1 UE case, due to lower power of distortion from the spatially filtered IMPs reaching UEs, as demonstrated in Table 5.3.

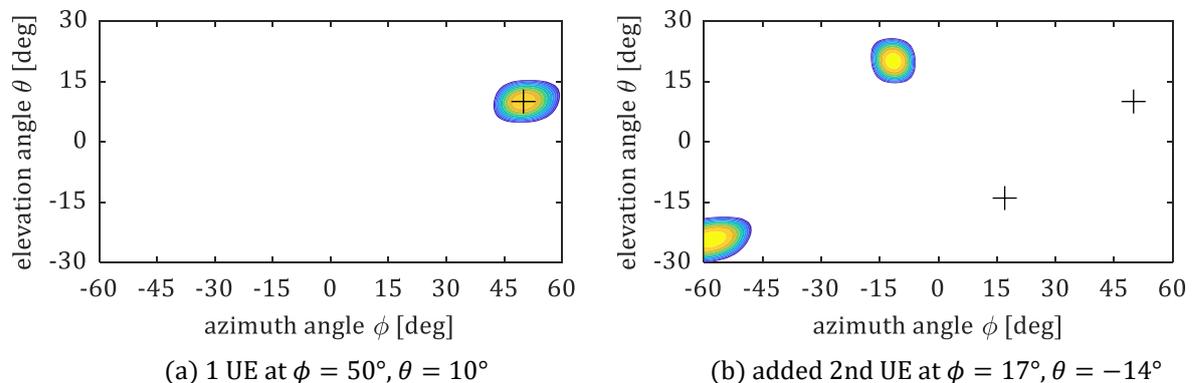


Figure 5.12: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses.

Next, using the 2 UE scenario from Figure 5.12(b) as a starting point, additional UEs were added one at a time. Apart from the 4th user, all the other UEs were added specifically by selecting angular direction with a relatively high power of spatially filtered IMPs. The intention was not to obtain a worst-case performance—which is expected to be on a par with the performance in the 1 UE case, as explained in Section 4.4.2.3—but instead provide an example of a relatively poor choice of UEs for spatially multiplexing in the same slot. The results of each step of this procedure are visualized with radiation patterns in Figure 5.13.

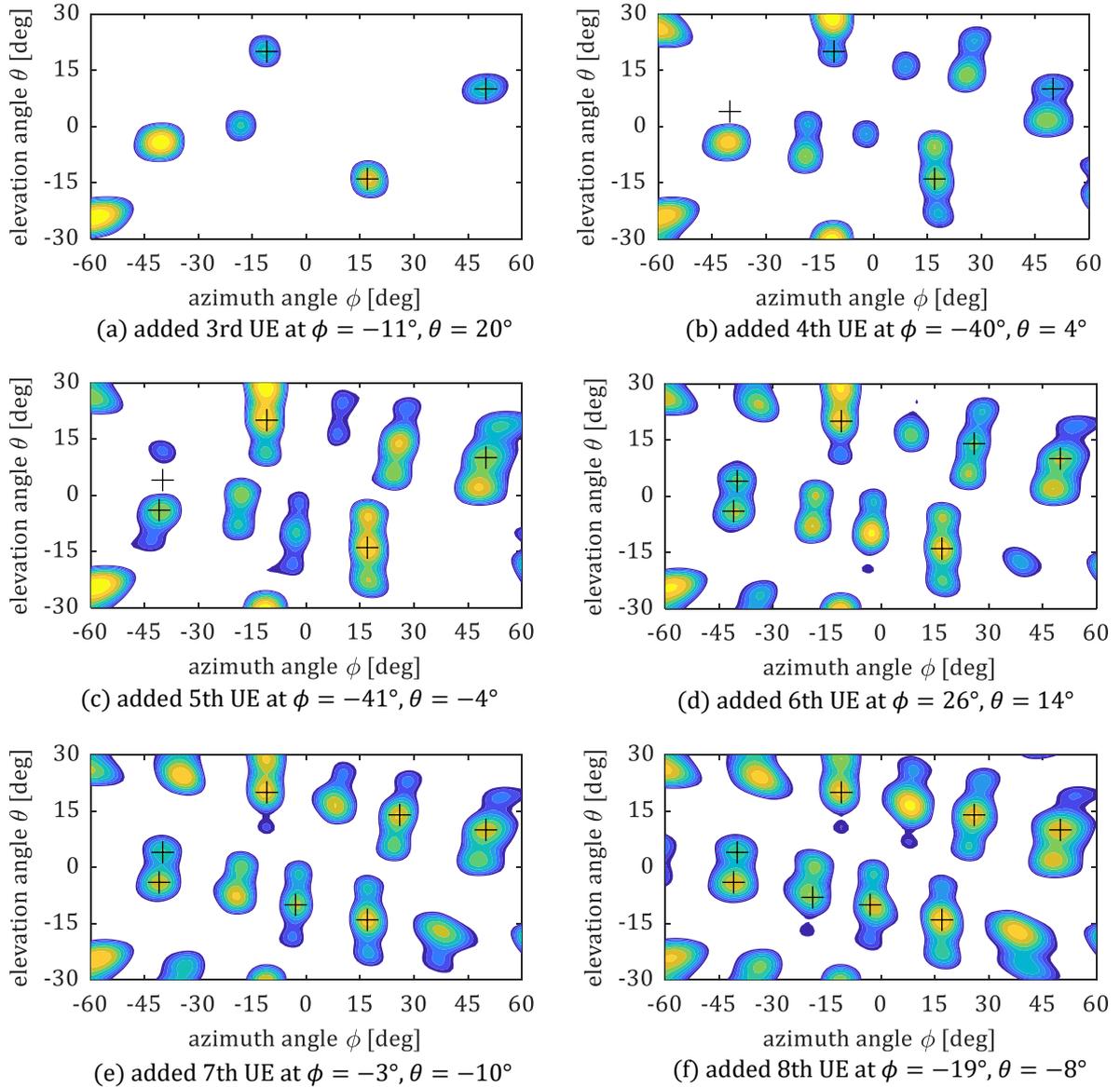


Figure 5.13: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses.

The mean values of the effective per UE array gains (applicable to the wanted data signal and the spatially filtered third-order IMPs), obtained by the precoding-based EVM performance prediction model, are presented in Table 5.3. Note that the array gains

applicable to the data signal were very close to the theoretical maximum achievable array gain of $10 \log_{10}(M/K)$, with up to about 0.03 dB lower effective array gain (resulting from the operation of the ZF precoding). When comparing the effective array gains applicable to the wanted and distortion signals, it seems clear that a noticeable portion of the spatially filtered IMPs can be expected to be received by UEs in this case, resulting in some level of EVM performance degradation compared to a typical case.

In order to provide a quantifiable example of the negative effects of receiving a portion of the spatially filtered IMPs, a 10 slot long transmission over free-space LoS channels was simulated in link-level with ICF (using a 6 dB peak reduction threshold) as the source of nonlinear distortion, for each of the scenarios presented in Figure 5.12 and Figure 5.13. The statistics of per UE SDRs (i.e., minimum, mean, and maximum) acquired from the LLS (calculated based on the per UE EVM performance results) are summarized in Table 5.3. Despite increasing the number of UEs, resulting in less explicitly user-directed IMPs (as explained in Section 4.4.2), relatively small SDR performance differences for the range from 4 to 8 UEs can be observed in Table 5.3.

Table 5.3: Summary of effective array gains and statistics of signal-to-distortion ratios (SDRs) for the UEs located at the directions from Figure 5.12 and Figure 5.13.

Number of UEs (K)	precoding-based EVM performance model		link-level simulation		
	mean of per UE array gains for the wanted data signal [dB]	mean of per UE array gains for the spatially filtered IMPs [dB]	min of SDRs [dB]	mean of SDRs [dB]	max of SDRs [dB]
1	21.1	21.1	23.5	23.5	23.5
2	18.1	-16.2	25.2	25.2	25.2
3	16.3	11.5	25.4	25.8	26.0
4	15.1	5.7	26.9	27.3	27.7
5	14.1	7.1	26.9	27.5	28.2
6	13.3	6.6	27.5	27.7	27.9
7	12.6	6.7	27.4	27.7	28.0
8	12.0	6.4	27.2	27.7	28.4

If the knowledge of the way the spatially filtered IMPs are expected to radiate is available in the RRM, distortion-aware decisions can be made. One obvious way would be to avoid extending the set of scheduled UEs by users that are at the angular directions of relatively high power from the spatially filtered IMPs of the already selected users. This may be feasible, provided that the scheduler has the option to delay traffic of the skipped users. An alternative would be to adjust the used modulation order and code rate, adapting to the expected SDR performance of a given scheduled set of UEs.

An example of a different selection of UEs for co-scheduling, using the 2 UE scenario from Figure 5.12(b) as a starting point, avoiding extending the scheduling set with UEs located at angles of a relatively high power of spatially filtered IMPs is considered next, to provide a performance comparison point to the scheduling decisions made in the example visualized in Figure 5.13. Instead of scheduling the 3rd UE located at $\phi = -11^\circ$ and $\theta = 20$, the 4th UE from Figure 5.13 was scheduled this time, followed by the 5th UE. The other 4 UEs in this new distortion-aware scheduling set are different than in the distortion-unaware one. In order to obtain a scenario with a relatively low level of the spatially filtered IMPs received by the UEs, the UEs from 5th to 8th were selected specifically from areas of angles with relatively low radiation of distortion, as visualized in Figure 5.14.

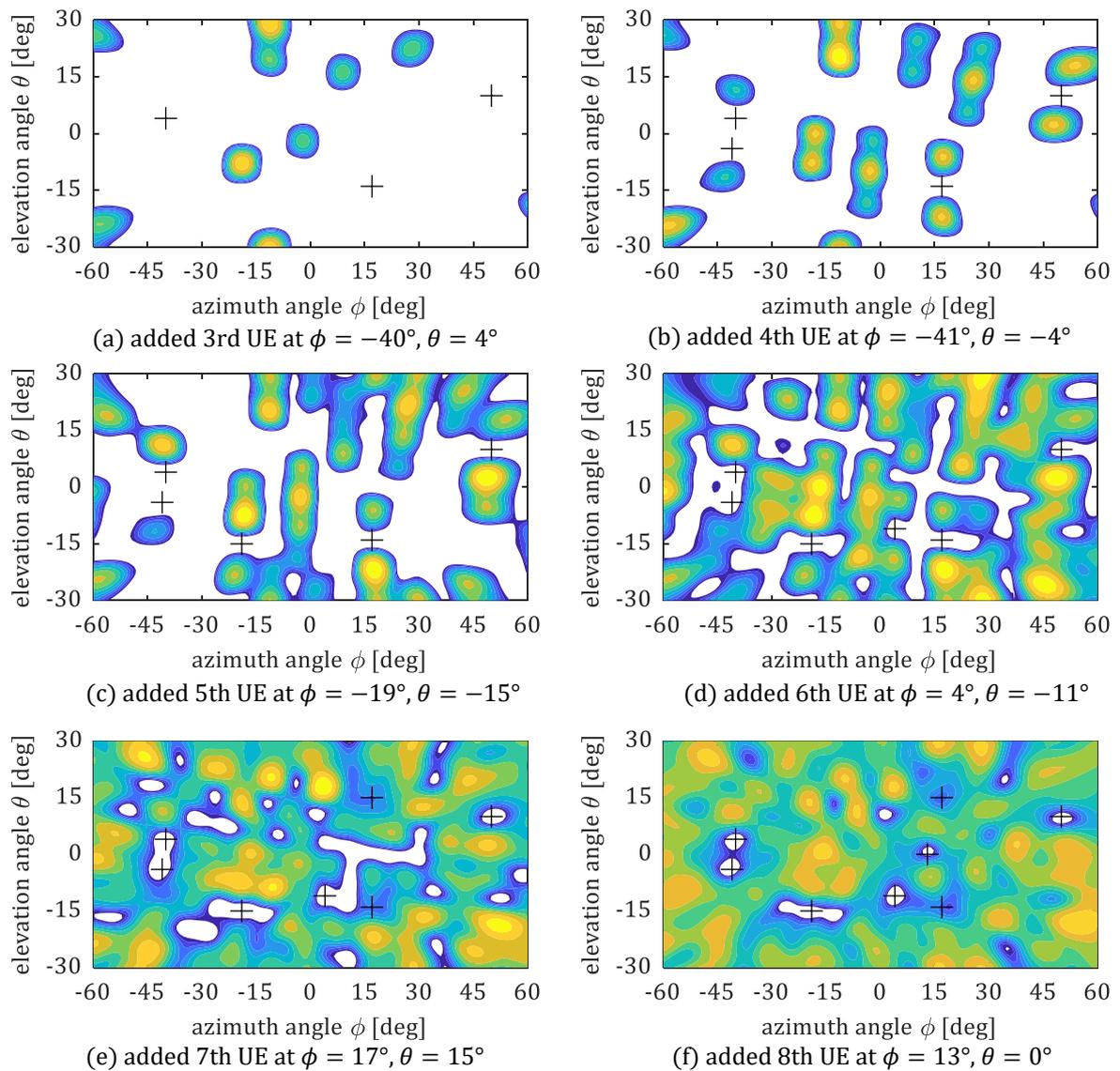


Figure 5.14: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses.

To quantify the performance impact of the distortion-aware scheduling operation, which effects are summarized in Figure 5.14, the same approach was used to obtain estimates of the effective array gains and statistics of SDRs as in case of the distortion-unaware scheduling example. When comparing the performance obtained for the relatively poor selection of UEs for co-scheduling in the distortion-unaware example, provided in Table 5.3, with the performance obtained in the distortion-aware example, summarized in Table 5.4 (only for the range of UEs from 3 to 8, since the performance for 1 and 2 UEs was not changed), clear differences in both the effective array gains and the SDRs can be observed. As expected, the effective array gain of the spatially filtered IMPs was significantly lower in the distortion-aware scheduling example (between about 12 dB and 29 dB lower), while the array gains applicable to the wanted signal were on par. That difference, resulting from the modified angular distribution of UEs, led to improved SDR performance. In this distortion-aware example, an improvement of SDR with increasing the number of UEs is visible in Table 5.4. This led to better SDR performance compared to the distortion-unaware example, for all of the presented SDR metrics in all of the verified scheduling cases. Up to about 3 dB better mean of SDRs was observed in the case of 8 UEs, when comparing with the results shared in Table 5.3.

Table 5.4: Summary of effective array gains and statistics of signal-to-distortion ratios (SDRs) for the UEs located and the directions from Figure 5.14.

Number of UEs (K)	precoding-based EVM performance model		link-level simulation		
	mean of per UE array gains for the wanted data signal [dB]	mean of per UE array gains for the spatially filtered IMPs [dB]	min of SDRs [dB]	mean of SDRs [dB]	max of SDRs [dB]
3	16.3	-17.4	26.8	26.9	26.9
4	15.0	-6.4	27.7	27.8	27.9
5	14.1	-8.8	28.4	28.6	28.8
6	13.3	-9.0	29.3	29.4	29.5
7	12.6	-6.2	29.8	30.1	30.3
8	12.0	-7.0	30.3	30.5	30.6

Clearly, it was proven that the decision on which UEs to spatially multiplex may impact the SDRs. This results from the relation between the radiation pattern of distortion (which depends on the array geometry and is determined by the used precoding weights) and the channel responses. The final impact on the achievable data rates depends on the SINR conditions the UEs are in.

Assuming equal SINR for all the co-scheduled UEs, the limiting effect of distortion in the user scheduling examples from Figure 5.13 and Figure 5.14 is considered next. To

calculate the impact on the per user data rates, the detailed per UE SDRs (which statistics are presented in Table 5.3 and Table 5.4) were used directly. This time, the limitation of the SE resulting from the maximum modulation order and code rate was ignored, to clearly show the potential differences between the distortion-unaware and distortion-aware decisions made in the considered examples of user scheduling.

A clear improvement in the achievable sum data rates per RE can be observed when the performance in the examples of distortion-unaware and distortion-aware scheduling examples, visualized in Figure 5.15, is compared. The largest observable data rate performance gap was obtained with 8 UEs (about 10% increase with the distortion-aware scheduling example in the SINR range from 40 to 50 dB), in line with the expectation based on the differences in the SDR performance.

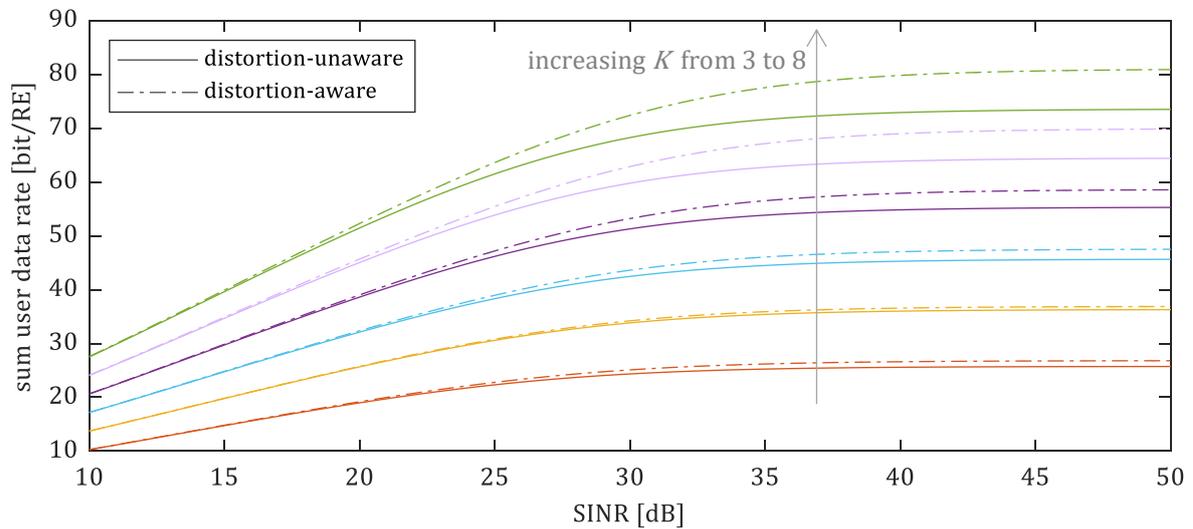


Figure 5.15: Comparison of the achievable per resource element (RE) sum data rates with the number of UEs going from 3 to 8, for distortion-unaware scheduling from Figure 5.13 and distortion-aware from Figure 5.14.

The two scheduling examples, visualized in Figure 5.13 and Figure 5.14, were provided to demonstrate that in principle an improvement in the achievable data rates (enabled by making distortion-aware user scheduling decisions based on the prediction of the radiation characteristic of distortion) is possible. It is expected that the practically attainable improvement in the data rate performance will not only vary depending on the channel characteristics of the UEs which need to be scheduled but will also depend on the capabilities and constraints of the employed distortion-aware scheduling and link adaptation solution.

5.4. Summary and Proof of Thesis

This chapter explored the transition from the analysis of the impact of residual TX impairments in Massive MIMO base stations focused on the signal transmission quality, covered in Chapter 4, to the system-level aspect of the interference- and noise-limited channels in a multi-cell network. By including the estimates of the received distortion power in the calculation of the SE, the impact on the achievable user data rates and cell throughput was investigated.

First, it was demonstrated in Section 5.3.1 that the channel hardening leads to comparable BLER performance for spatially correlated frequency-selective LoS and NLoS channels, based on CDL-C and CDL-D channel models with 100 ns RMS DS. With increasing the number of antennas, the SNR required for 10% BLER with CDL-C (NLoS) was reducing to finally reach the same BLER performance as in CDL-D (LoS). Each increase of the number of antennas provided smaller improvement, with most of the 6.8 dB reduction in the required SNR observed when going from 1 to 8 antennas already. The comparison of the BLER performance obtained by separately analyzing the impact of RX noise and TX distortion suggests that the impact of noise and fully correlated distortion is similar. It is expected that the difference between the required SNR and SDR of about 0.8 dB observed in the results may be influenced by the strategy of power allocation over SCs and the equalization performance.

Additionally, the results shared in Section 5.3.1 confirmed that the SNDR required to reach a given BLER (translating to achievable SE) is few decibels higher than based on the channel capacity calculation, which assumes a Gaussian input signal.

Next, the analysis of the impact of unequal power allocation between users, covered in terms of EVM performance in Section 4.7.5, was extended in Section 5.3.2 to capture the impact on the data rate performance, considering the limiting influence of the RX noise. It was not only demonstrated that the impact of distortion is more severe when UEs are in better SNR conditions, but more importantly that the strategy of power allocation for equal SNR fails to achieve the target of comparable QoS when the limiting effect of distortion is considered.

Shifting the focus from an isolated single cell to a multi-cell network, the impacts of intra- and inter-cell interference were covered using numerical simulations in Section 5.3.3, reflecting the SINR distribution in a 16-cell network with single cell and multi-cell precoding algorithms and resulting limits on the SE. By introducing the expected power of received

distortion, estimated based on the best-case EVM performance prediction model from (4.19) into the calculation of the SE, the combined effects of interference, noise, and distortion were observed. It was demonstrated that the SINR distribution in the network noticeably improved when the number of antennas was increased, leading to more users that could potentially become distortion-limited. Additionally, it was observed that the performance gap between the M-MMSE precoding (i.e., being able to deal with both the intra- and inter-cell interference) and RZF precoding (i.e., covering the intra-cell interference), represented by the distance between the SINR distributions, was increasing with increasing number of antennas. It is worth to point out that negative effects of distortion were relatively small with per antenna EVM of about -30 dB and less. This results both from the SINR distribution (with less than about 10% of UE transmissions having SINR higher than 30 dB, in the range between 10 to 100 antennas that was verified with distortion) and from the expectation that most of the distortion radiates away from users when transmitting to 10 UEs. However, when the per antenna EVM was increased, a visible degradation of the achievable data rates was observed (i.e., about 6% for -20 dB, 24% for -10 dB, and 60% for 0 dB). This proves that the effects of the residual TX impairments cannot be generally ignored, even when the number of antennas goes to infinity, since the portion of the energy of nonlinear distortion reaching UEs depends on the number of UEs that the base station is transmitting to.

Next, the cell throughput was linked with the energy consumption of the base station in Section 5.3.3, by relating the IBO with the SDR and making simplified assumptions on the dependence of the energy efficiency operating point on the IBO. It was demonstrated that usage of PAs with higher peak power capabilities (resulting in larger IBO) leads to increase in average cell throughput at the cost of increased power usage. However, when the IBO was high enough to result in negligible levels of received distortion (compared to the levels of received interference and noise), further reduction of EVM resulting from increasing IBO led to increase in the energy consumption without noticeable positive impact on the cell throughput performance. This suggests that depending on the SINR distribution expected in a given network deployment (influenced not only by the physical locations of UEs, but also by the array geometry and the utilized precoding schemes) there is a certain level of tolerable per antenna EVM that leads to marginal impact on the user data rates and cell throughput, enabling consideration of a balance between the acceptable impact on the data rates and the energy consumption of the base station.

Finally, the potential benefits of using estimates of the power of received distortion as input for user scheduling decisions were demonstrated in Section 5.3.4. The impact of

distortion explored based on the best-case EVM performance may be considered representative for the typical performance. However, in the worst-case the nonlinear distortion may fully radiate towards the UEs, as explained based on the EVM performance in Section 4.4.2.3, potentially causing a substantial drop in the instantaneous data rates. It was demonstrated in Section 5.3.4 for an example of scheduling 8 UEs that a relatively poor selection of spatially multiplexed UEs, made specifically based on prediction of radiation pattern of spatially filtered IMPs for a rectangular array, may result in degraded SDR performance, compared to a situation that co-scheduling certain UEs was avoided, based on the expected increase in the portion of received distortion. The gap in the sum data rate between the selected examples of distortion-unaware and distortion-aware scheduling was increasing when the improved SINR conditions of UEs were assumed, with the largest gap of about 10% observed for 8 UEs in the SINR range from 40 to 50 dB.

While the two scheduling examples were provided to demonstrate that in principle an improvement in the achievable data rates is possible, validating the second thesis of this dissertation, it is expected that the practically attainable improvement in the data rate performance will not only vary depending on the per antenna EVM level and the channel characteristics but will also depend on the capabilities and constraints of the employed distortion-aware scheduling and link adaptation solution.

Chapter 6

Exploiting Spatial Degrees of Freedom to Compensate Transmitter Distortion

The distortion introduced by the PAPR reduction may have negative influence on the achievable DL data rates and energy consumption of Massive MIMO base stations, as demonstrated in Chapters 4 and 5. While there is a class of PAPR reduction solutions that avoid radiating in-band distortion towards UEs, resulting in a distortion-free transmission, the SINR conditions in practical deployments allow the UEs to tolerate certain level of error (resulting from both distortion and IUI) with negligible impact on the performance of the data transmission. Hence, the target of a distortion-free transmission seems to be over-constraining the PAPR reduction problem, leading to unnecessarily increased energy consumption, resulting from the additional computational complexity spent in the digital signal processing part of the TX chain to address this problem. Therefore, the potential for improving the performance of a low-complexity conventional per antenna PAPR reduction solution based on the signal distortion is explored here.

The variation of the average power over antennas, highlighted in [2], can degrade the BER performance of a per antenna limiter-based PAPR reduction, as demonstrated in [48]. One strategy for dealing with this situation is to reduce the per antenna power variation. For example, by exploiting the degrees of freedom available in the Massive MIMO channel to select different precoding weights than resulting directly from the pseudoinverse of the channel matrix, like in the FA-based beamforming proposed in [48]. An alternative strategy, explored here, is to adjust the PAPR reduction approach to exploit the unequal distribution of the average power over antennas, to improve the performance of the data transmission. The proposed solution addresses the distortion introduced in a subset of antennas by transmitting compensation signals from another subset of antennas. However, instead of reserving antennas for compensation, as in AR-based approaches [56] – [60], the power headroom in a subset of antennas is utilized to provide that compensation. Different sources of distortion, other than resulting from PAPR reduction, could also be considered to be handled by the same solution framework, similarly to the proposed extension of AR-based approach to handle PA nonlinearities in [58].

This chapter covers a solution that was described originally in a patent application (by the PhD candidate and the assistant supervisor), which was not published yet at the time of writing of this dissertation (as of August 2024). Part of the content of this chapter is based on that material.

6.1. System Model

The system model of transmission from a Massive MIMO base station with M antennas to K single-antenna UEs defined in Section 4.1 is followed here. In the context of that model, the proposed compensation method for distortion introduced by the ICF-based PAPR reduction is described in Section 6.3.

6.2. Preliminaries

While there are no antennas reserved for transmitting compensation signals in the proposed solution, the fundamental principles of using the excess degrees of freedom offered by having more TX antennas than UEs to compensate for distortion introduced in a subset of antennas are similar to the AR-based PAPR reduction approach, which is briefly introduced in Section 6.2.1.

The proposed solution exploits the variation of the per antenna power caused by precoding to improve the EVM performance of a conventional per antenna PAPR reduction approach. An extreme case of such a variation of the per antenna power can be experienced with tapered antenna arrays, which are introduced in Section 6.2.2.

6.2.1. Using Reserved Antennas to Compensate Transmitter Distortion

The idea to reserve a subset of antennas in the base station for transmitting solely the compensation signals designed to correct for the distortion introduced in the main antennas (transmitting the data signals) was proposed in [56] and explored further in [57] – [60], focusing primarily on compensating the error introduced by the conventional per antenna PAPR reduction solutions based on signal distortion, like companding transforms, CF-, or PC-based approaches.

Generally, only a partial compensation of distortion could be achieved using the AR-based method, even if a perfect CSI knowledge is available at the base station side. This is related with the distortion of the compensation signal in the reserved antennas, resulting either from the nonlinear PAs as in [56] – [60], or an explicit PAPR reduction as in [73] (co-authored by the PhD candidate). The compensation signal has much lower average power

and significantly higher PAPR than the data signal, as demonstrated in Figure 6.1 for the case of a 6 dB peak power threshold and compensation with either 16 or 8 out of 64 antennas.

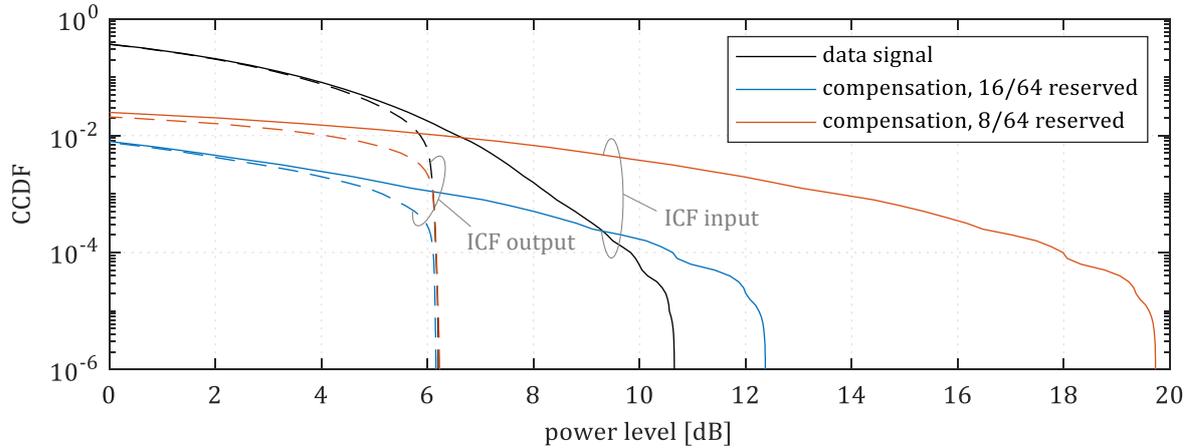


Figure 6.1: Comparison of the complementary cumulative distribution functions (CCDFs) of the instantaneous power of the data and compensation signals for either 16 or 8 out of 64 antennas reserved, at the input and output of the iterative clipping and filtering (ICF) block. Based on a figure originally published in [73] (co-authored by the PhD candidate).

The issue of distorting the compensation signals is especially pronounced when the per antenna output power is increased (i.e., the IBO is reduced, leading to operating PAs closer to their saturation levels) in the main antennas to compensate for the apparent reduction of both the array gain and the sum output power, caused by dedicating a subset of antennas for transmitting solely the compensation signal. For the optimal reservation level of about 20–25% antennas, suggested in [58] – [59], this translates to about 2.0–2.5 dB loss in the SINR (resulting from the reduction of the power of the data signal) that needs to be handled, by increasing the power in the main antennas. This leads either to more distortion generated in the main antennas, which degrades the EVM performance, or to the need to use PAs with a higher peak power capability at a larger IBO, which results in increased energy consumption.

6.2.2. Per Antenna Power in a Tapered Array

Tapering of the antenna elements in the array is a method to influence the shape of the radiation pattern by using a nonuniform amplitude distribution [74, Sec. 6.8], with a gradually reducing per antenna power from the center elements towards the edges of the antenna array. This allows to control the level of inter-beam interference, by reducing the power in the side lobes of the radiation pattern, when a precoding without handling IUI is used, like the MR precoding.

When tapering is used, the level of side lobes is reduced at the cost of widening the main beam and a reduction of directivity, as visualized in Figure 6.2 for tapering weights based

on Hamming and Hann windows compared with no tapering for a linear array with 32 omni-directional antennas. The per antenna output power for the case without tapering was 30 dBm, leading to 45 dBm EIRP in the UE direction resulting from the 15 dB array gain associated with 32 antennas. While the same sum output power was used in all the cases, the reduction of the power in the peak of the main lobe (pointing to the 30° azimuth angle, using beam-steering), visible in Figure 6.2 for the radiation patterns of the tapered arrays, was 1.4 dB for tapering weights based on both Hamming window and Hann window (generated from a window length of 34, ignoring the 2 outermost zero-valued weights).

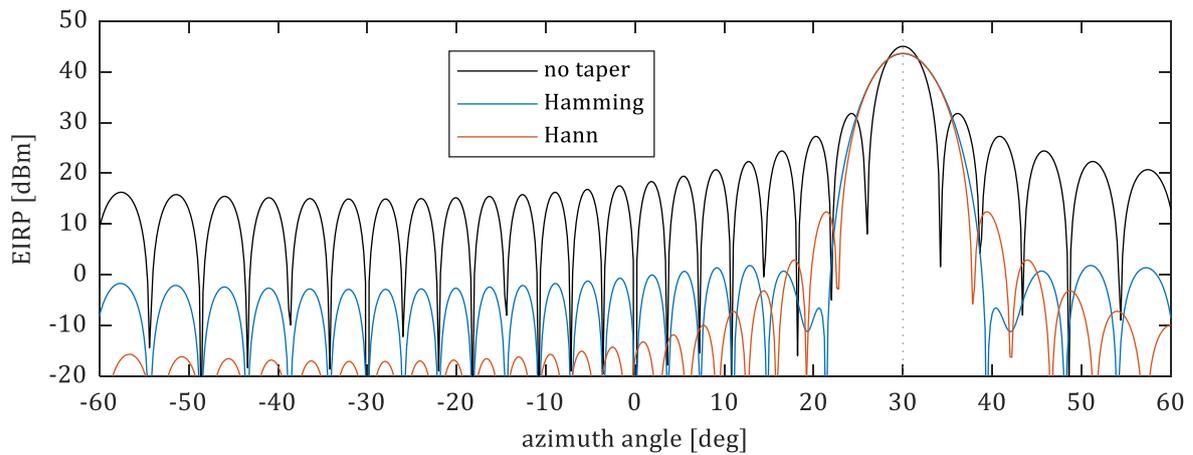


Figure 6.2: Comparison of radiation patterns for a linear array of 32 omni-directional antennas, with a beam-steering to 30° azimuth angle for no tapering, as well as tapering based on Hamming and Hann windows.

The usage of tapering inevitably introduces a variation of the expected power levels over antennas. If a PAPC was enforced, the gain of the per antenna tapering weight would be limited to the level resulting from an equal distribution of the sum output power over antennas, which is considered here to be the nominal power level. However, when the same sum output power as in the case of no tapering is targeted, there are some antennas in which the expected power level is increased above the nominal level (by tapering weights having gain higher than 0 dB), as demonstrated in Figure 6.3(a) for a linear array of 32 antennas. In this case, the gain range over antennas is from about -17.8 dB to about 4.1 dB for the tapering weights based on the Hamming window and from about -36.8 dB to about 4.1 dB for the Hann window. At the same time, majority of antennas are having a power headroom (gain of the tapering weight below 0 dB), as visualized in Figure 6.3(b) for linear arrays of 8, 32, and 1024 antennas.

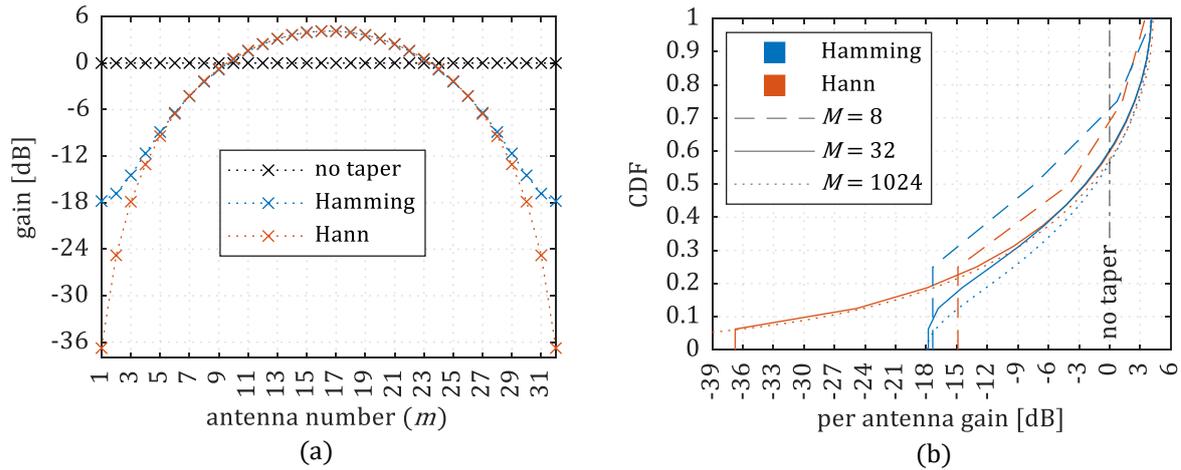


Figure 6.3: A comparison of the gains of tapering weights based on Hamming and Hann windows: (a) per antenna gain for a linear array of 32 antennas; (b) cumulative distribution functions (CDFs) of the per antenna gains, depending on the number of antennas M in a linear array.

Assuming that PAs with the same saturation levels are used in all the antenna paths, tapering leads to operating one subset of antennas with a relatively large IBO from the peak power threshold, reducing the energy efficiency for these antennas, while having another subset of antennas exceed the nominal level, operating with a smaller IBO (higher energy efficiency) but with a degraded EVM (due to an effective reduction of CR, leading to higher EVM as demonstrated in Section 4.7.1). That per antenna EVM may be not reduced to negligible levels when considering the OTA EVM performance, due to partially coherently combining at the UE location, as explained in Section 4.4.2.

6.3. Proposed Compensation Method for Signal Distortion Introduced by Peak-to-Average-Power Ratio Reduction

The excess degrees of freedom in the Massive MIMO channel, available thanks to having more transmit antennas M than the UEs K , allow to improve the handling of the PAPR reduction compared with the conventional per antenna PAPR reduction approaches [39]. Among different techniques using these spatial degrees of freedom, listed in Section 3.3, reservation of a subset of antennas for transmitting solely a compensation signal, introduced in [56], is a starting point for the proposed novel solution.

The downside of an AR-based compensation solution is the apparent reduction of the array gain and widening of the lobes in the radiation pattern of the array, as explained in Section 6.2.1. The proposed solution does not rely on reserving antennas, but instead hinges on splitting the M antennas of the array into 3 disjoint sets: distorting \mathcal{D} , compensating \mathcal{C}

(having a power headroom), and untended \mathcal{U} (where these untended antennas may be either antennas with a power headroom or without it). The distortion introduced in $M_{\mathcal{D}}$ distorting antennas is partially compensated for in $M_{\mathcal{C}}$ compensating antennas, by combining the compensation signals with the data signals before jointly going through a part of TX processing chain which can distort them. Hence, generally the processing chains of antennas in sets \mathcal{C} and \mathcal{U} can introduce distortion, but that part of distortion is not compensated for.

Allocation of antennas to these sets is done based on the headroom between the per antenna mean power operating point (or a peak power capability or both) and the nominal level of the mean power. Hence in the end, the proposed technique is based on using a power headroom, so it is in a way a power headroom reservation technique. While this reservation may be done explicitly on purpose, or resulting from existing power headroom introduced by tapering (as covered in Section 6.2.2), or resulting from application of precoding that does not explicitly consider a PAPC, leading to variation of the expected power level over antennas that changes in the scheduling rhythm (here 1 slot).

Simplifying the definition of the RX signal for single-cell operation from (4.4), assuming that the IUIs are handled by precoding and ignoring the RX noise, the signal received by the k -th UE on the j -th SC can be described as

$$y_k = \underbrace{\mathbf{h}_k^T \mathbf{w}_k s_k}_{\text{Wanted signal}} + \underbrace{\mathbf{h}_k^T \mathbf{d}}_{\text{Transmitter distortion}} \quad (6.1)$$

where $y_k \in \mathbb{C}$ is a frequency-domain data sample received by the k -th UE, $\mathbf{h}_k^T \in \mathbb{C}^{1 \times M}$ represents channel responses from M base station's antennas to the UE antenna, $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the k -th UE precoding vectors multiplied by the data sample s_k , and $\mathbf{d} \in \mathbb{C}^{M \times 1}$ is a vector representing TX distortion.

The effectively received unwanted distortion from (6.1) can be defined as

$$u_k = \mathbf{h}_k^T \mathbf{d}. \quad (6.2)$$

The proposed solution modifies the data transmitted from the antennas in \mathcal{C} , with the target to reduce u_k . Clearly, a full compensation would be achieved if the compensation symbol received on the j -th SC was equal to $-u_k$. However, this may not be achievable in practice with the proposed solution, as explained in Section 6.3.1.

6.3.1. Compensating Distortion from a Subset of Antennas

The proposed approach for compensating distortion addresses only a portion of the distortion from (6.2), associated with the distorting antennas set \mathcal{D} , as visualized in Figure 6.4. Therefore, a full compensation is not expected, since there can be remaining uncompensated

distortion of the wanted signal introduced in the subset of compensating antennas \mathcal{C} . Moreover, similarly to the issue highlighted for AR-based approaches in Section 6.2.1, the compensation signals may get distorted together with the wanted signal in the PAPR reduction blocks of the compensating antennas, visualized in Figure 6.4.

However, in practice a full compensation is not really required, since certain level of EVM can be tolerated by UEs, as pointed out in Section 4.3.1. Therefore, it should be enough to provide a partial compensation to improve the performance of the PAPR reduction solution in terms of EVM (and BLER as a result) in comparison to using a conventional per antenna PAPR reduction without any compensation of the introduced distortion.

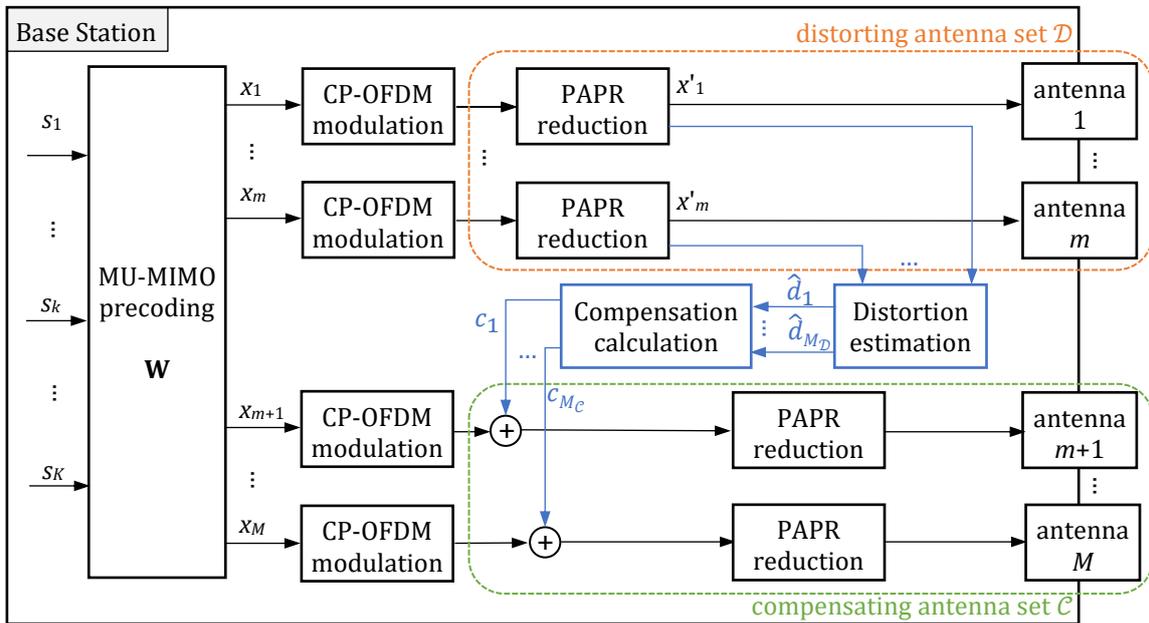


Figure 6.4: Block diagram of the proposed compensation approach using antennas with a power headroom. Only selected elements of the processing chain, relevant from the perspective of the proposed method, shown.

Note that while the block diagram in Figure 6.4 shows antennas from 1 to m allocated to the distorting set \mathcal{D} , from $m+1$ to M allocated to the compensating set \mathcal{C} , and no antennas in the unintended set \mathcal{U} , the allocations of antennas to the sets can be generally different (including potentially dependence on the array geometry, similarly to considerations in [60] for AR-based PAPR reduction). While the processing after CP-OFDM modulation block is performed in the time domain, a frequency-domain representation is used here for brevity.

If the compensation signal was transmitted without any alteration in the compensating antennas set \mathcal{C} , the signal received by the k -th UE on the j -th SC would be

$$y_k = \underbrace{\mathbf{h}_k^T \mathbf{w}_k s_k}_{\text{Wanted signal}} + \underbrace{\mathbf{h}_k^T \mathbf{d}}_{\text{Transmitter distortion}} + \underbrace{\mathbf{h}_k^{cT} \mathbf{c}}_{\text{Distortion compensation}} \quad (6.3)$$

where $\mathbf{h}_k^{cT} \in \mathbb{C}^{1 \times M_c}$ represents channel responses from M_c base station's antennas (covering all the antennas in the compensating set \mathcal{C}), to the UE antenna, and $\mathbf{c} \in \mathbb{C}^{M_c \times 1}$ is a vector representing the distortion compensation signals. However, it is expected that the compensation signal \mathbf{c} added to the data signal \mathbf{x} in the compensating antennas will be subject to distortion. Therefore, the signal received by the k -th UE on the j -th SC can be generally defined as

$$y_k = \underbrace{\mathbf{h}_k^T \mathbf{w}_k s_k}_{\text{Wanted signal}} + \underbrace{\mathbf{h}_k^T \mathbf{d}}_{\substack{\text{Distortion} \\ \text{of the wanted signal}}} + \underbrace{\mathbf{h}_k^{cT} (\mathbf{c} + \boldsymbol{\delta})}_{\text{Distorted compensation signal}} \quad (6.4)$$

where $\boldsymbol{\delta} \in \mathbb{C}^{M_c \times 1}$ is the distortion of the compensation signal introduced in the transmission chains of the antennas in the compensating subset \mathcal{C} .

Note that the elements of \mathbf{h}_k^{cT} belong to a subset of the elements of \mathbf{h}_k^T . Hence, there is no additional overhead related with acquiring CSI for the reserved antennas, since all the antennas are used for transmission and reception of the data signals. This is a fundamental difference compared to an AR-based compensation, where the elements of the channel vector for the reserved antennas (transmitting the compensation signal) are not a subset of the channel vector for the primary antennas (transmitting the distorted data signal).

The only relevant source of distortion here is assumed to be the PAPR reduction. For a per antenna operating limiter-based PAPR reduction, d_m is the signal that can be added to the input of the PAPR reduction block x_m in the m -th antenna processing chain in order to achieve a reduction of the peak power, as visualized for ICF in Figure 4.4. In the simplified diagram of the DL processing chain presented in Figure 6.4, $x'_m = x_m + d_m$ is the signal that is transmitted for the m -th antenna in the distorting set \mathcal{D} . Assuming the PAPR-reducing distortion signal d_m is explicitly available directly at the output of the PAPR reduction block, it can be simply forwarded for all the $M_{\mathcal{D}}$ antennas in the distorting set \mathcal{D} to the block that calculates the compensation signals for all the M_c antennas in the compensating set \mathcal{C} . Otherwise, the block handling the distortion estimation visualized in Figure 6.4 needs to separately calculate the distortion estimation signals \hat{d}_m , before passing them to the block handling calculation of compensation signals, described in more detail in Section 6.3.2.

6.3.2. Calculation of the Compensation Signals

One of possible approaches to calculate the per antenna compensation signals c_m is outlined in Figure 6.5, from the perspective of a frequency-domain processing. Any potentially required pre-processing steps, like down-conversion, down-sampling, and CP-OFDM demodulation of the time-domain distortion estimation signals before the input of that block, as well as any required post-processing steps like CP-OFDM modulation of the compensation signal after the output of that block are omitted here for brevity.

For the j -th SC, the distortion estimates for the distorting antennas $\hat{\mathbf{d}}^D \in \mathbb{C}^{M_D \times 1}$ are used first to calculate estimates of the expected effectively received distortion $\hat{u}_k^D \in \mathbb{C}$ for each of the UEs, using the channel estimates $\mathbf{h}_k^{D^T} \in \mathbb{C}^{1 \times M_D}$.

Based on these per SC per UE estimates, compensation symbols a_k can be calculated and then precoded for the set of compensating antennas, using a precoding matrix $\mathbf{V} \in \mathbb{C}^{M_c \times K}$ calculated based on the channel estimates (equal to the per UE channel responses $\mathbf{h}_k^{c^T} \in \mathbb{C}^{1 \times M_c}$, when a perfect CSI knowledge on the base station side is assumed). The precoding approach for the compensation signal is not directly dependent on the method used for the wanted signal. Hence, for example, MR instead of ZF precoding could be considered to be used to reduce the complexity of the compensation calculation block.

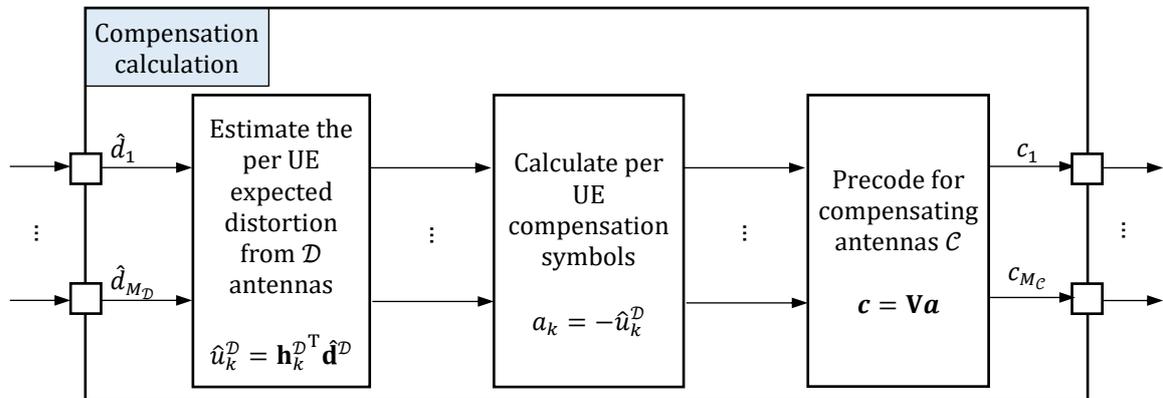


Figure 6.5: Block diagram of the compensation calculation block, from Figure 6.4, for a single subcarrier.

6.4. Methods

The methods defined in Section 4.6, for the fundamental study of the way the distortion introduced by the residual TX impairments behave in Massive MIMO channels, are followed here as well.

The simulation results, obtained with the proprietary LLS developed by the PhD candidate extended to cover the proposed approach for compensation of the PAPR reduction

distortion described in Section 6.3, are provided for linear array of antennas tapered with weights based on the Hamming window. The compensation signal is precoded using ZF in most of the cases, apart from the specific verification of the feasibility of using low-complexity MR precoding, where it is explicitly explained.

Two sets of randomly selected azimuth angles were generated for the MU-MIMO simulations, with a minimum angular separation of UEs wider than the constraint used for a non-tapered array in Section 5.3, to accommodate for a wider main lobe experienced with a tapered array.

6.5. Simulation Results and Discussion

With LLS, it is demonstrated that the proposed approach leads to better performance compared to using the conventional per antenna ICF.

First, the efficacy of the proposed solution in compensating distortion is demonstrated for a single UE transmission case in Section 6.5.1 from the perspective of the peak power capability required to achieve a given fixed EVM target. Next, the improvement in the EVM performance is demonstrated for MU-MIMO scenarios in Section 6.5.2, for transmission in both free-space LoS and multi-path channels, for a given fixed level of the peak power capability of the PAs. Finally, the potential impact on the energy efficiency and energy consumption is explored in Section 6.5.3, based on a selected MU-MIMO transmission case.

6.5.1. Reduction of the Required Peak Power Capability

One of the potential benefits of being able to reduce a portion of the distortion that is effectively received by the UE is to reduce the required IBO, leading to a lower level of the peak power capability required for maintaining a given level of the average output power.

In order to demonstrate that potential, transmission from a linear array of 32 omnidirectional antennas to a single UE over a free-space LoS channel was verified. The peak power capability required to not exceed EVM of -28 dB while simultaneously achieving the same level of the mean received power at the UE side as for a non-tapered array was identified, assuming PAs in all the antennas using the same saturation levels. Note that this power target means that the sum output power of a tapered array was larger than for the non-tapered array (in order to additionally compensate for the reduction in directivity, which is a more ambitious target than the target of compensating for the reduction in the sum output power used in all the other simulation results in this chapter).

The identified minimum required peak power levels, for a number of compensating antennas M_c from 0 to 16, is presented in Figure 6.6. By using all of the antennas that operate with a power headroom (i.e., here 16 out of the total of 32) for compensation, the target was achieved without an increase of the peak power capability compared with the non-tapered array case.

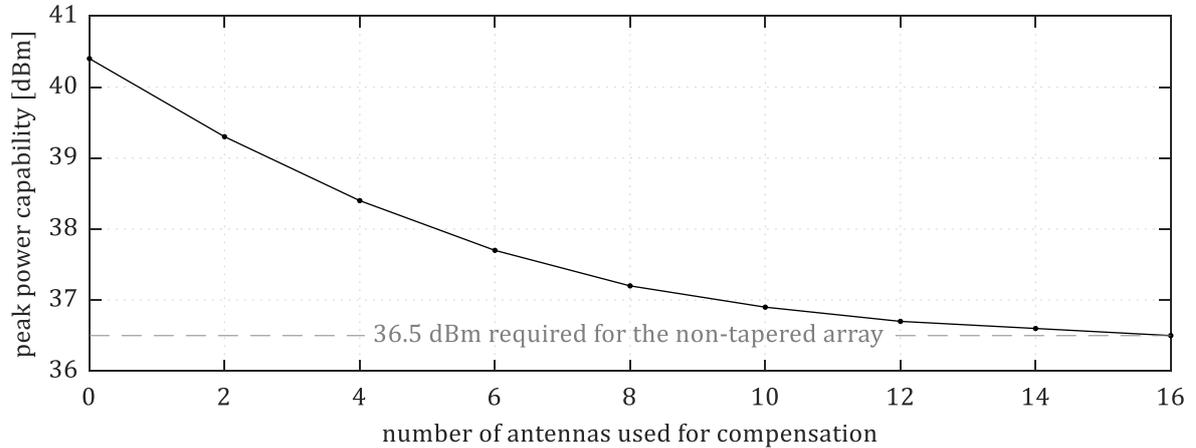


Figure 6.6: The peak power capability required for an array of 32 antennas using Hamming-based tapering to achieve the same error level and mean power at the user as for a non-tapered array reduces with increasing the number of compensating antennas M_c .

If the compensation of the PAPR reduction distortion was not applied, but both the peak power capabilities and the RX power targets were kept like for the non-tapered array, the EVM would be -15.4 dB instead of the targeted -28.0 dB. Which means that the power of the effectively received distortion was substantially reduced when the compensation was utilized (i.e., by about 12 dB with the same IBO as in the non-tapered array case when 50% of antennas were used for compensation). This proves that a noticeable EVM performance improvement could be expected, in at least a subset of scenarios.

As visible in Figure 6.6, it was also possible to use only a subset of the antennas that operated with a power headroom and still achieve a substantial reduction of the required peak power capabilities. Even using only 2 antennas (i.e., less than 7% of all the antennas) reduced the required PAs' peak power capabilities by more than 1 dB, while using 8 antennas brought a reduction by more than 3 dB (i.e., from 40.4 dBm down to 37.2 dBm). This shows that a trade-off between the added complexity in the digital signal processing part of the DL chain and the energy consumption of the analog part of the processing chain can be made.

While the general efficacy of the proposed distortion compensation method was demonstrated here (in a transmission to a single UE over a free-space LoS channel) for the use case of reducing the peak power capability required for a fixed EVM target, the EVM

performance improvement potential for a fixed peak power threshold is considered in the next section (with a focus on the MU-MIMO transmission scenarios).

6.5.2. Improvement in the Level of Error at the Users

To demonstrate the potential impact on the DL throughput, the EVM performance for a sample of MU-MIMO scenarios with randomly selected angular locations of UEs was verified in simulations. Starting with the performance for a free-space LoS transmission, covering both ZF and MR precoding of the compensation signals, followed by the results for transmission over frequency-selective LoS and NLoS channels.

The only notable source of distortion here was the PAPR reduction realized by the per antenna ICF, configured with a CR of 6 dB. This means that the peak power limitation level was 6 dB above the nominal power, leading to different effective levels of PAPR in different antennas of the tapered linear array (using a total of 32 antennas).

The simulation results summarized in Table 6.1, present RMS of EVM calculated over UEs for a 1 slot transmission over a free-space LoS channel. On the one hand, it is visible that the EVM for the tapered array without compensation significantly degraded in comparison to the non-tapered array reference, which was expected due to a larger portion of the total energy being effectively subject to a lower (i.e., stricter) PAPR constraint. For that case, the degradation was increasing (from 5.2 dB up to 8.0 dB) when the number of UEs was increasing. On the other hand, it is clearly visible based on the results presented in Table 6.1 that usage of the proposed compensation approach led to a noticeably improved EVM performance. Even usage of a little bit more than 10% of antennas for compensation led to noticeable EVM improvements in relation to the not compensated case, especially for 1 and 2 UEs. The results with 50% of antennas used for compensation demonstrate a substantial improvement potential, with even more than 16 dB EVM reduction for the scenarios with 4 and 6 UEs.

Table 6.1: Comparison of EVMs with and without compensation of transmitter distortion.

Number of UEs (K)	Angular directions of UEs [deg]	RMS of EVM over UEs [dB]				
		non-tapered reference	tapered, with different ratio of antennas belonging to the compensating set			
			0 / 32	4 / 32	8 / 32	16 / 32
1	37.8	-24.1	-18.9	-23.0	-27.9	-30.5
2	37.8, -44.8	-25.6	-19.7	-25.2	-31.9	-34.0
4	37.8, -44.8, 15.9, -26.6	-27.4	-19.5	-22.0	-34.7	-36.9
6	37.8, -44.8, 15.9, -26.6, 54.9, -1.8	-28.2	-20.2	-20.8	-30.6	-36.9

Note that the power levels of the signals received by the UEs were different for non-tapered, tapered without compensation, and tapered with compensation cases covered in Table 6.1. The signal levels for the case of tapered with compensation were about 1.4 dB lower than for the non-tapered reference, which matches with the expected reduction in the array gain for the array with 32 antennas using tapering weights based on the Hamming window, as explained in Section 6.2.2. However, it is worth noting that the cases without compensation saw a loss between about 1.7 dB and 2.1 dB. Which means that, in practice, the compensation improved not only the distortion level at the UE (visible in the EVM) but also the signal level at the UE, improving the SINR as a result.

While it is likely that the level of improvement depends on the angular locations of UEs, for example due to the way the IM beams radiate and how much of the compensation signal targeted for other UEs will be received by different UEs, as explored in Section 4.4.2 for the cases without compensation. However, the simulation results presented in Table 6.1 demonstrate that a significant improvement in EVM performance could be expected, at least in a subset of spatial scenarios.

The EVM performance improvement potential is clear when judging the RMS of EVM over UEs, presented in Table 6.1. However, it is worth noting that different levels of EVM improvement were observed between UEs. An increased variation of EVM over UEs, in relation to the case without compensation, was considered to be worth exploring. The summary of the statistics of EVM over UEs for the scenarios from Table 6.1 is presented in Figure 6.7 for both ZF and MR precoding of the compensation signal, with 8 antennas in the compensating set \mathcal{C} . The proposed solution clearly improved the median of EVMs (visible as the horizontal bars within the bodies of the boxplots in Figure 6.7) compared with the tapered uncompensated case, for both ZF and MR precoding of the compensation signal. While a relatively large variation of EVM over UEs can be observed, especially for MR (notably with 2 UEs in the 6 UEs case experiencing up to 2 dB degradation in EVM compared to the case without compensation), the usage of ZF precoding consistently resulted in a significant EVM improvement for each of the UEs.

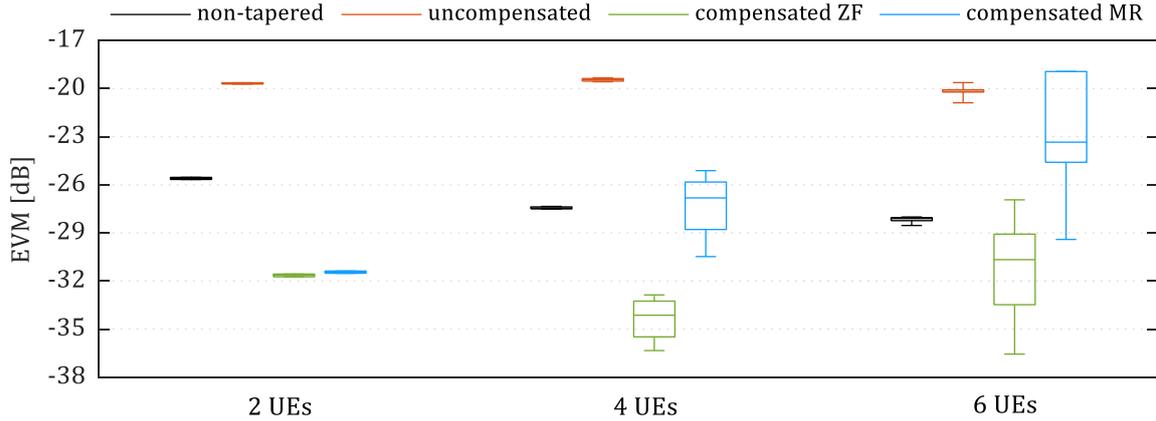


Figure 6.7: Boxplots representing the statistics of EVM over UEs, for the scenarios from Table 6.1, achieved with a non-tapered array as well as a tapered array without compensation and with compensation, using zero forcing (ZF) and maximum ratio (MR) precoding of the compensation signals.

In order to demonstrate the feasibility of using the proposed compensation solution in frequency-selective scenarios, a set of randomly generated angular locations for up to 6 UEs was explored for transmission over spatially correlated CDL-D (LoS) and CDL-A (NLoS) channels (with 100 ns RMS DS) for a tapered linear array of 32 antennas. The summary of RMS EVM over UEs for a single slot simulation with no compensation and with compensation in 50% antennas is presented in Table 6.2. For the multi-path LoS channel, the improvement in EVM performance of about 10 dB can be observed consistently for all the verified scenarios. While the EVM improvements for 4 and 6 UE cases for the NLoS channel were much lower than for the channel with LoS, as visible in Table 6.2, it is worth noting that a significant variation of EVM over UEs was observed in these cases (with more than 10 dB, with EVM about -30 dB for at least one of the UEs in each of these scenarios).

Table 6.2: EVMs for a tapered array with 0 or 16 antennas out of 32 in the compensating set \mathcal{C} .

Number of UEs (K)	Angular directions of UEs [deg]	RMS EVM over UEs [dB]			
		CDL-D (LoS)		CDL-A (NLoS)	
		$M_c = 0$	$M_c = 16$	$M_c = 0$	$M_c = 16$
1	-15.1	-19.0	-30.1	-22.4	-31.7
2	-15.1, 54.1	-20.2	-30.2	-21.6	-34.3
4	-15.1, 54.1, 27.8, 11.8	-19.6	-29.9	-17.5	-21.9
6	-15.1, 54.1, 27.8, 11.8, -41.3, -57.5	-19.4	-29.9	-18.3	-22.1

Clearly, the EVM performance can be improved using the proposed method of compensation, as demonstrated for both free-space LoS and multi-path LoS and NLoS scenarios. Next, the aspects related with the energy consumption are considered, by exploring in detail the per antenna power distribution in the 6 UE scenario from Table 6.2.

6.5.3. Impact on the Energy Consumption

On the one hand, utilization of the proposed compensation method to enable usage of PAs with a lower peak power capability in the antennas of the distorting set \mathcal{D} , explored in Section 6.5.1, allows to use smaller IBOs and operate at higher energy efficiency operating points (i.e., closer to the PAs's saturation). On the other hand, the compensating antennas operate with higher levels of average power, compared to not using the power headroom for compensation, and with relatively large PAPRs in case of tapered arrays, leading to increased peak power capability requirement for the antennas in the compensating set \mathcal{C} .

To demonstrate the impact on the energy consumption, the per antenna power measurement results for the simulation case of 6 UEs for CDL-D (LoS) channel from Table 6.2 were analyzed. The potential differences in the energy consumption come from the per antenna average power level and its relation to the peak power capability, impacting the energy efficiency.

The variation of both the average and peak power levels for a tapered antenna array, without and with compensation (using 8 antennas at each of the edges of the array, out of 32 available), is presented in Figure 6.8. It is visible that the peak power level at the ICF input has a similar distribution over antennas as the average power, which is the result of applying tapering weights based on the Hamming window, however there is some additional variation resulting from the relatively short length of the simulation of 1 slot. With the peak power threshold configured to be 6 dB above the nominal level of the average power there is a substantial reduction of the peak power observed at the ICF output for the antennas closer to the middle of the array. This power reduction results also in decreasing the average power level in the antennas in the middle, visible by the distance between the ICF input and output in Figure 6.8.

The difference in the power distribution at the ICF output between the case without compensation and with compensation is clearly visible for both peak and average powers in Figure 6.8. The proposed PAPR reduction scheme utilizes the power headroom available in the 8 outermost antennas at each of the edges of the array to transmit a distorted compensation signal together with the data signal. This is visible in the apparent increase of both peak and average power levels for antennas from 1st to 8th and from 25th to 32nd. It is worth noting that the power budget was not optimally exploited, since the peak power level for the 2 antennas at each of the edges was below the 6 dB threshold, suggesting that more

improvement in the EVM performance could be achieved with a different power allocation of the compensation signal to the antennas than the equal distribution used here.

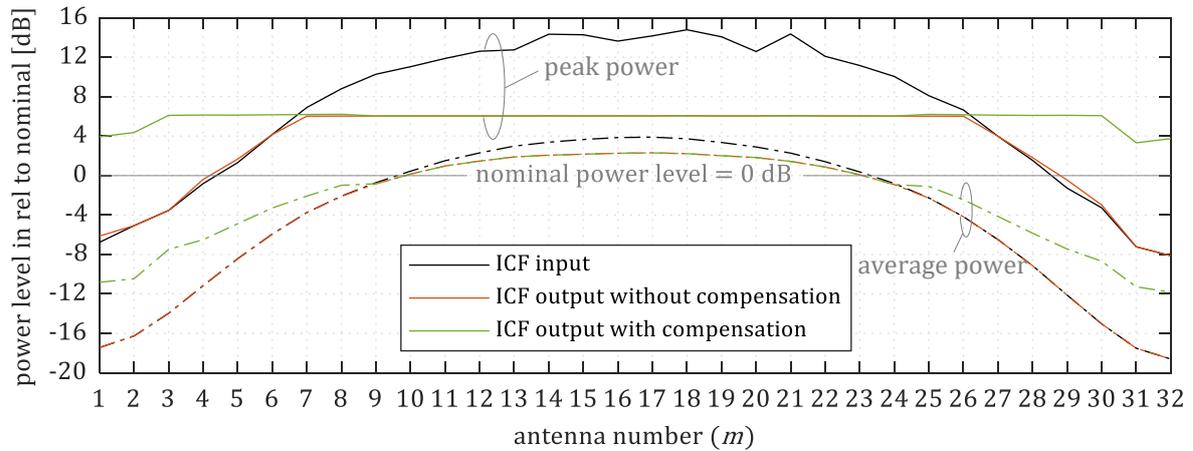


Figure 6.8: Comparison of the per antenna distribution of the peak and average power levels at the input and output of the iterative clipping and filtering (ICF) block for tapered antenna array without and with compensation of the PAPR reduction distortion (in 50% of the antennas, located symmetrically at the edges).

In order to fairly compare the performance of the proposed compensation solution, additional simulations without compensation were performed to achieve similar EVM performance operating points while maintaining comparable level of received power, to obtain equivalent SINR performance. Since an increased IBO was required in this case, PAs with higher peak power capabilities for the uncompensated case were required. In case of using PAs with the same peak power limit in all the antennas, about 4.5 dB increase in the peak power capability (i.e., to 10.5 dB PAPR) was required. The alternative option of using PAs with different peak power limits (but the same PAPR limits), led to the PAPR constraint of about 7.5 dB (resulting in reduction of the peak power limits in many antennas, but increase of that limit in the middle antennas up to about 11.5 dB of PAPR).

Using a similar approach for estimating the difference in the energy consumption as in [73], assuming 30 dBm mean output power and 40% energy efficiency operating point for a PA with 6 dB IBO from its peak power capability (which seems to be a reasonable assumption, since the state-of-the-art GaN-based PAs could operate at about 40% efficiency with 7 dB IBO [7]) of 36 dBm, and for simplicity a 1 percentage points change in the PA efficiency for a 1 dB change in the IBO the energy consumption of PAs can be calculated. With a non-tapered array following these assumptions a power supply of 80 W would be required to provide the sum output power of 32 W (i.e., 1 W per antenna).

The per antenna energy efficiency and the resulting power usage is presented in Figure 6.9 for the cases of tapered linear array: with compensation (with equal peak power limit over all antennas); uncompensated, with equal peak power limit; uncompensated, with equal PAPR limit. For tapered arrays, the overall energy consumption is primarily impacted by the antennas closer to the middle of the array (which transmit most of the overall energy, as visible in Figure 6.8), hence the energy efficiency operating points of these antennas matter the most. Clearly, the energy efficiency levels for the compensated case are better than for both of the uncompensated cases in majority of the antennas closer to the middle (as visible in Figure 6.9(a), for all the antennas when compared to the case of equal peak limit and for antennas from 8th to 25th when compared to the equal PAPR limit). At the same time, the lower power usage for the case with compensation observable for the middle antennas in Figure 6.9(b) results additionally from a reduced output power in these antennas (caused by more energy removed from the wanted signals due to effectively lower IBO, compared to the uncompensated cases) achieved at the cost of increasing the output power in the antennas of the compensating set, resulting in the higher power usage observable for antennas from 1st to 8th and from 25th to 32nd in Figure 6.9(b).

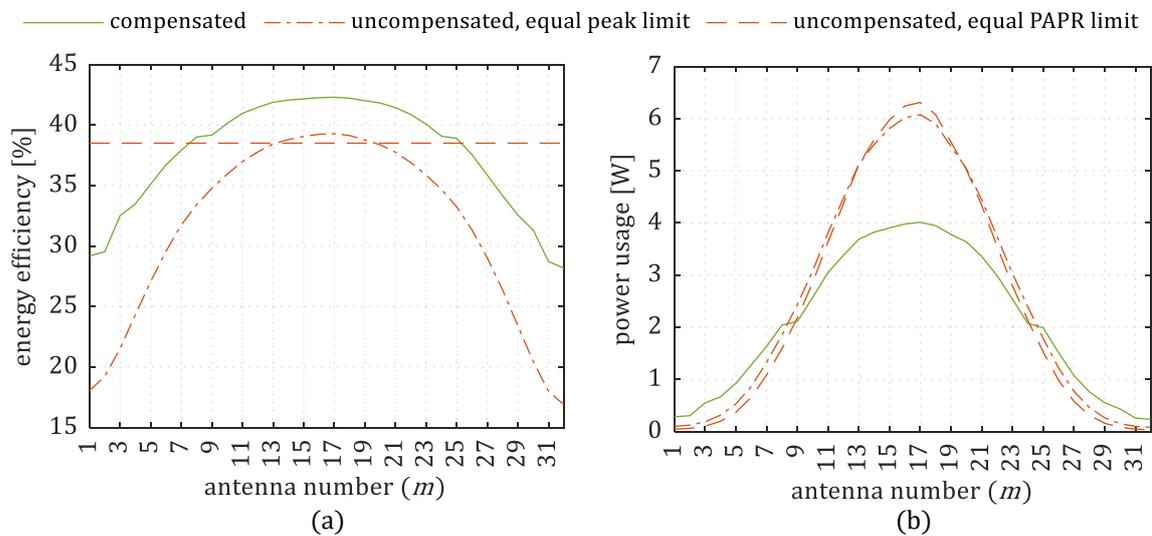


Figure 6.9: Comparison of the performance for tapered antenna array, in the 6 UE CDL-D (LoS) channel case from Table 6.2, for compensated and uncompensated array (with either equal peak or PAPR limits, over antennas): (a) energy efficiency; (b) power usage.

The sum power usage of the PAs in the tapered array, based on the per antenna usage presented in Figure 6.9(b), was about: 67 W for the compensated case; 82 W for the uncompensated case with equal peak limit; 80 W for the equal PAPR limit case. Surely, different results will be obtained for different transmission scenarios (including a potential

dependence on the number of UEs, their angular locations, and channel characteristics) and with different assumptions on the scaling of energy efficiency. However, the reduction in the energy consumption (of more than 15%), demonstrated with the simulation results for a single MU-MIMO scenario in Figure 6.9, provides an evidence that an improvement in the energy consumption in the PAs could be expected, at least in a subset of spatial scenarios. This suggests that the proposed solution has the potential of improving the balance between the EVM performance and the overall energy consumption of a Massive MIMO base station, provided that the additional energy consumed in the digital signal processing part of the TX chain (required to provide the distortion compensation functionality) is lower than the reduction achieved in the analog part of the TX chain.

6.6. Summary and Proof of Thesis

A novel approach for exploiting the excess degrees of freedom available in the Massive MIMO channel, by using the power headroom in a subset of antennas to compensate for the distortion introduced in a different subset of antennas, was proposed and explored here.

Using the ICF-based PAPR reduction as the source of distortion in a tapered linear array, the efficacy of the proposed solution was demonstrated in LLS. First, the feasibility of reducing the required IBO for a fixed EVM and SINR targets was proven for transmission to a single UE over a free-space LoS channel, with the same peak power capability as in case of non-tapered array achieved when 50% of antennas were allocated to the compensating set. Next, the improvement potential of the EVM performance was exemplified for MU-MIMO transmission over free-space LoS, multi-path LoS, and NLoS channels. Significant improvements were visible, especially for LoS scenarios, with increasing the number of antennas allocated to the compensating set. Finally, the potential for reducing the energy consumption in the analog part of the TX chain was demonstrated.

The third thesis of this dissertation, expressing the feasibility of partially compensating for the distortion introduced by a limiter-based PAPR reduction in a subset of antennas using a different subset of antennas with a power headroom, was validated here for the proposed compensation solution based on the set of simulations covering the use case of tapered antenna arrays using the conventional per antenna ICF-based PAPR reduction approach.

Chapter 7

Conclusions

The ultimate objective of the research effort covered by this dissertation was to answer whether the residual transmitter (TX) impairments introduced due to non-idealities in the hardware (HW) of practical Massive MIMO base stations, with compact arrays of co-located antennas, may have noticeable negative impact on the achievable downlink (DL) data rates.

The analytical analysis of the radiation characteristic of nonlinear distortion introduced either by saturation in the power amplifiers (PAs) or by peak-to-average-power ratio (PAPR) reduction based on signal distortion was supported by the results of the link-level simulation (LLS). Additionally, the HW impairments that could be relatively accurately represented by additive Gaussian noise uncorrelated between antennas of the array (which may potentially represent the effects of quantization) were covered to some extent.

A special attention was given to Massive MIMO base stations with a moderate number of antennas (e.g., 32), as used in initial 5G deployments, and to the differences resulting from the array geometry, especially the angular granularity achievable with not so massive number of columns (in the range from 8 to 32) in a rectangular array.

First, the primary thesis of this dissertation, expressing the dependence of the error level at the user on both the structure of the base station and the selection of precoding weights, was explored. The performance of a 5G Massive MIMO base station was characterized in terms of the error vector magnitude (EVM), as seen from the user equipment (UE) perspective. It was shown analytically and demonstrated in simulation that the nonlinear distortion can be expected to combine, at least to some extent, coherently at the UE locations. In fact, it was proven that when precoding matrix is aligned with the transmission over deterministic free-space line-of-sight (LoS) channels there are worst case combinations of UE angular directions that result in the nonlinear distortion radiating only in the UE directions (benefiting from the full array gain) for any number antennas M and users K —for both linear and rectangular antenna arrays.

It was also demonstrated in numerical simulation, using Monte Carlo trials of the UE angular locations, that a spectrum of EVM performance levels could be expected, depending on the M and K . At the same time, it was shown that typically the EVM could be expected to be closer to the best-case scenario of UE locations and the probability of the worst-case

was relatively low, resulting from the apparent heavy-tailed distribution of the EVM over randomly selected sets of UE angular locations.

The expectations for the EVM in transmission over free-space LoS channels based on the relation between the radiation patterns of the wanted data signal and nonlinear distortion, calculated based on the channel and precoding matrices, based on the intermodulation products (IMPs) expected to be introduced by a third-order nonlinearity, were compared with the results of the LLS for the PAPR reduction based on iterative clipping and filtering (ICF) as well as linearized PAs. Based on the comparison of the EVM performance results obtained for deterministic free-space LoS, as well as spatially correlated frequency-selective LoS and NLoS channels, with the results of numerical simulations, it was concluded that the EVM performance could be predicted using a statistical model based on third-order IMPs with acceptable accuracy.

Additionally, the impact of the number of UEs and the power distribution between them on the EVM performance with quantization noise was considered. The simulation results suggest that the impact of distortion should be considered when selecting the power allocation strategy, to avoid significantly degrading the performance of UEs by utilizing large differences between the powers of different transmission layers.

Next, based on the foundation given by the characterized EVM performance, the impact on the achievable user data rates and cell throughput in DL was considered. The achievable spectral efficiency (SE), limited not only by the channel capacity, but also by the maximum modulation order and code rate supported in 5G New Radio transmission, formed a basis for calculating the limits for the achievable DL data rates. The calculation of the SE limit was extended by reflecting the received power of distortion, in addition to the signal-to-interference-plus-noise ratio (SINR) representing the channel quality perceived by the UE.

The user data rates and cell throughput for different PA power operating points (resulting from the peak power capabilities and the applied input back-off), leading to different EVM and energy efficiency operating points, was considered. A potential for improvement of the balance between the achievable DL data rates and energy consumption by using the EVM expectations in distortion-aware scheduling was demonstrated, validating the second thesis of this dissertation.

Finally, a novel approach for exploiting the excess degrees of freedom available in the Massive MIMO channel, hinging on using the power headroom in a subset of antennas to compensate for the distortion introduced in a different subset of antennas, was proposed. The third thesis of this dissertation, expressing the feasibility of partially compensating for the

distortion introduced by a limiter-based PAPR reduction in a subset of antennas using a different subset of antennas with a power headroom, was validated for the use case of tapered antenna arrays using the conventional per antenna ICF-based PAPR reduction approach. A significant EVM performance improvement potential was demonstrated in the LLS, especially for scenarios with LoS (including with frequency-selective fading). The potential for reducing the energy consumption in the analog part of the TX chain was also demonstrated.

One of the future research directions, basing on the foundations addressed in this dissertation, could be focused on improving the accuracy and expanding the scope of the proposed EVM performance prediction models. On the one hand side, an extension of these statistical models by reflecting the impact of higher-order IMPs could be considered, to improve the EVM prediction accuracy. On the other hand, it may be more consequential to delve into expanding the scope of the EVM prediction to cover also the isolator-free TX designs, by addressing the impacts of the mutual coupling between array elements as well as the impedance mismatches between PAs and antennas. At the same time, characterization of the impact on the EVM performance of the interactions between the conventional compact arrays of co-located antennas, antennas deployed in a distributed fashion, and reflecting intelligent surfaces—in the use cases of both cellular and user-centric cell-free Massive MIMO networks—could be considered an exciting future research project.

Bibliography

- [1] E. Björnson, J. Hoydis and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3–4, pp. 154–655, 2017, doi: 10.1561/20000000093.
- [2] C. Mollén, U. Gustavsson, T. Eriksson and E. G. Larsson, "Spatial characteristics of distortion radiated from antenna arrays with transceiver nonlinearities," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, p. 6663–6679, Oct. 2018, doi: 10.1109/TWC.2018.2861872.
- [3] International Telecommunication Union (ITU), "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation ITU-R M.2083-0, 2015. Accessed: Sep. 12, 2024. [Online]. Available: <https://www.itu.int/rec/R-REC-M.2083>.
- [4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [5] Next Generation Mobile Networks (NGMN) Alliance, "Network energy efficiency phase 2," Oct. 2023. Accessed: Sep. 12, 2024. [Online]. Available: <https://www.ngmn.org/publications/network-energy-efficiency-phase-2.html>.
- [6] 3GPP, "Study on network energy savings for NR (Release 18)," TR 38.864, Mar. 2023. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.864/38864-i10.zip.
- [7] S. Wesemann, J. Du and H. Viswanathan, "Energy efficient extreme MIMO: Design goals and directions," *IEEE Communications Magazine*, vol. 61, no. 10, p. 132–138, Oct. 2023, doi: 10.1109/MCOM.004.2200958.
- [8] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, p. 3590–3600, Nov. 2010, doi: 10.1109/TWC.2010.092810.091092.
- [9] T. L. Marzetta, E. G. Larsson, H. Yang and H. Q. Ngo, *Fundamentals of Massive MIMO*, Cambridge University Press, 2016, doi:10.1017/CBO9781316799895.

- [10] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, p. 186–195, Feb. 2014, doi: 10.1109/MCOM.2014.6736761.
- [11] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis and T. L. Marzetta, "Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, p. 3–20, 2019, doi:10.1016/j.dsp.2019.06.007.
- [12] E. Björnson, E. G. Larsson and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, p. 114–123, Feb. 2016, doi: 10.1109/MCOM.2016.7402270.
- [13] L. Sanguinetti, E. Björnson and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Transactions on Communications*, vol. 68, no. 1, p. 232–257, Jan. 2020, doi: 10.1109/TCOMM.2019.2945792.
- [14] E. Björnson, J. Hoydis and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, p. 574–590, Jan. 2018, doi: 10.1109/TWC.2017.2768423.
- [15] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, p. 40–60, Jan. 2013, doi: 10.1109/MSP.2011.2178495.
- [16] J. Hoydis, S. ten Brink and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, p. 160–171, Feb. 2013, doi: 10.1109/JSAC.2013.130205.
- [17] E. Björnson, J. Hoydis, M. Kountouris and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Transactions on Information Theory*, vol. 60, no. 11, p. 7112–7139, Nov. 2014, doi: 10.1109/TIT.2014.2354403.
- [18] E. Björnson, P. Zetterberg, M. Bengtsson and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE*

- Communications Letters*, vol. 17, no. 1, p. 91–94, Jan. 2013, doi: 10.1109/LCOMM.2012.112012.122003.
- [19] U. Gustavsson, C. Sánchez-Perez, T. Eriksson, F. Athley, G. Durisi, P. Landin, K. Hausmair, C. Fager and L. Svensson, "On the impact of hardware impairments on massive MIMO," in *2014 IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 294–300, doi: 10.1109/GLOCOMW.2014.7063447.
- [20] Y. Zou, O. Raeesi, L. Antilla, A. Hakkarainen, J. Vieira, F. Tufvesson, Q. Cui and M. Valkama, "Impact of power amplifier nonlinearities in multi-user massive MIMO downlink," in *2015 IEEE Globecom Workshops (GC Wkshps)*, 2015, pp. 1–7, doi: 10.1109/GLOCOMW.2015.7414011.
- [21] E. G. Larsson and L. Van der Perre, "Out-of-band radiation from antenna arrays clarified," *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 610–613, Aug. 2018, doi: 10.1109/LWC.2018.2802519.
- [22] W. Sandrin, "Spatial distribution of intermodulation products in active phased array antennas," *IEEE Transactions on Antennas and Propagation*, vol. 21, no. 6, p. 864–868, Nov. 1973, doi: 10.1109/TAP.1973.1140612.
- [23] C. Hemmi, "Pattern characteristics of harmonic and intermodulation products in broadband active transmit arrays," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 6, p. 858–865, June 2002, doi: 10.1109/TAP.2002.1017668.
- [24] L. Anttila, A. Brihuega and M. Valkama, "On antenna array out-of-band emissions," *IEEE Wireless Communications Letters*, vol. 8, no. 6, p. 1653–1656, Dec. 2019, doi: 10.1109/LWC.2019.2934442.
- [25] C. Mollen, E. G. Larsson, U. Gustavsson, T. Eriksson and R. W. Heath, "Out-of-band radiation from large antenna arrays," *IEEE Communications Magazine*, vol. 56, no. 4, p. 196–203, April 2018, doi: 10.1109/MCOM.2018.1601063.
- [26] C. Fager, X. Bland, K. Hausmair, J. Chani Cahuana and T. Eriksson, "Prediction of smart antenna transmitter characteristics using a new behavioral modeling approach," in *2014 IEEE MTT-S International Microwave Symposium (IMS2014)*, 2014, pp. 1–4, doi: 10.1109/MWSYM.2014.6848309.
- [27] K. Hausmair, S. Gustafsson, C. Sánchez-Pérez, P. N. Landin, U. Gustavsson, T. Eriksson and C. Fager, "Prediction of nonlinear distortion in wideband active

- antenna arrays," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 11, p. 4550–4563, Nov. 2017, doi: 10.1109/TMTT.2017.2699962.
- [28] C. Fager, T. Eriksson, F. Barradas, K. Hausmair, T. Cunha and J. C. Pedro, "Linearity and efficiency in 5G transmitters: New techniques for analyzing efficiency, linearity, and linearization in a 5G active antenna transmitter context," *IEEE Microwave Magazine*, vol. 20, no. 5, p. 35–49, May 2019, doi: 10.1109/MMM.2019.2898020.
- [29] L. Monteyne, G. Callebaut, B. Sihlbom and L. Van der Perre, "Out-of-band distortion in massive MIMO: What to expect under realistic conditions?," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, 2021, pp. 1–5, doi: 10.1109/VTC2021-Fall52928.2021.9625447.
- [30] F. Rottenberg, G. Callebaut and L. Van der Perre, "Spatial distribution of distortion due to nonlinear power amplification in distributed massive MIMO," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 76–80, doi: 10.1109/SPAWC51858.2021.9593246.
- [31] M. B. Salman, E. Björnson, G. M. Gùvensen and T. Çilođlu, "Analytical nonlinear distortion characterization for frequency-selective massive MIMO channels," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 6535–6540, doi: 10.1109/ICC45041.2023.10279061.
- [32] N. Kolomvakis and E. Björnson, "Nonlinear distortion issues created by active reconfigurable intelligent surfaces," in *2024 18th European Conference on Antennas and Propagation (EuCAP)*, 2024, pp. 1–5, doi: 10.23919/EuCAP60739.2024.10501268.
- [33] N. Kolomvakis, A. Kosasih and E. Björnson, "Nonlinear distortion radiated from large arrays and active reconfigurable intelligent surfaces," *arXiv*, doi: 10.48550/arXiv.2401.12622.
- [34] U. Gustavsson, E. Larsson and C. Mollén, "Power amplifier-aware user scheduling". United States Patent US 2019/0373616 A1, 5 Dec. 2019.
- [35] S. R. Aghdam, S. Jacobsson and T. Eriksson, "Distortion-aware linear precoding for millimeter-wave multiuser MISO downlink," in *2019 IEEE*

- International Conference on Communications Workshops (ICC Workshops)*, 2019, pp. 1–6, doi: 10.1109/ICCW.2019.8757031.
- [36] F. Rottenberg, G. Callebaut and L. Van der Perre, "Z3RO precoder canceling nonlinear power amplifier distortion in large array systems," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 432–437, doi: 10.1109/ICC45855.2022.9838378.
- [37] F. Rottenberg, G. Callebaut and L. Van der Perre, "The Z3RO family of precoders cancelling nonlinear power amplification distortion in large array systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 3, p. 2036–2047, March 2023, doi: 10.1109/TWC.2022.3208436.
- [38] T. Feys, G. Callebaut, L. Van der Perre and F. Rottenberg, "Measurement-Based Validation of Z3RO precoder to prevent nonlinear amplifier distortion in massive MIMO systems," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5, doi: 10.1109/VTC2022-Spring54318.2022.9861020.
- [39] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: A survey and taxonomy," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1567–1592, 4th Quart. 2013, doi: 10.1109/SURV.2013.021313.00164.
- [40] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, p. 303–313, Feb. 2013, doi: 10.1109/JSAC.2013.130217.
- [41] S. Taner and C. Studer, " ℓ_p - ℓ_q -norm minimization for joint precoding and peak-to-average-power ratio reduction," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 437–442, doi: 10.1109/IEEECONF53345.2021.9723175.
- [42] L. Hua, Y. Wang, Z. Lian, Y. Su and Z. Xie, "Low-complexity PAPR-aware precoding for massive MIMO-OFDM downlink systems," *IEEE Wireless Communications Letters*, vol. 11, no. 7, p. 1339–1343, Jul. 2022, doi: 10.1109/LWC.2022.3166892.
- [43] L. Hua, Y. Wang, Z. Lian, Y. Su and Z. Xie, "LADMM-based PAPR-aware precoding for massive MIMO-OFDM downlink systems," *IEEE Transactions on*

- Vehicular Technology*, vol. 72, no. 1, p. 735–746, Jan. 2023, doi: 10.1109/TVT.2022.3203859.
- [44] S. Taner and C. Studer, "Alternating projections method for joint precoding and peak-to-average-power ratio reduction," in *2023 IEEE Wireless Communications and Networking Conference (WCNC)*, 2023, pp. 1–7, doi: 10.1109/WCNC55385.2023.10118766.
- [45] H. Bao, J. Fang, Q. Wan, Z. Chen and T. Jiang, "An ADMM approach for PAPR reduction for large-scale MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, p. 7407–7418, Aug. 2018, doi: 10.1109/TVT.2018.2837112.
- [46] R. Zayani, H. Shaiek and D. Roviras, "PAPR-aware massive MIMO-OFDM downlink," *IEEE Access*, vol. 7, p. 25474–25484, 2019, doi: 10.1109/ACCESS.2019.2900128.
- [47] Y. Matsumoto, K. Tateishi and K. Higuchi, "Performance evaluations on adaptive PAPR reduction method using null space in MIMO channel for eigenmode massive MIMO-OFDM signals," in *2017 23rd Asia-Pacific Conference on Communications (APCC)*, 2017, pp. 1–6, doi: 10.23919/APCC.2017.8304036.
- [48] M. Suzuki, Y. Kishiyama and K. Higuchi, "Combination of beamforming with per-antenna power constraint and adaptive PAPR reduction method using null space in MIMO channel for multiuser massive MIMO-OFDM transmission," in *2018 21st International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2018, pp. 134–139, doi: 10.1109/WPMC.2018.8712912.
- [49] T. Kageyama, O. Muta and H. Gacanin, "An adaptive peak cancellation method for linear-precoded MIMO-OFDM signals," in *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2015, pp. 271–275, doi: 10.1109/PIMRC.2015.7343308.
- [50] T. Suzuki, M. Suzuki, Y. Kishiyama and K. Higuchi, "Complexity-reduced adaptive PAPR reduction method using null space in MIMO channel for MIMO-OFDM signals," *IEICE Transactions on Communications*, vol. E103.B, no. 9, p. 1019–1029, 2020, 10.1587/transcom.2019EBT0005.

- [51] J. Saito, T. Hara and K. Higuchi, "Efficient parallel PAPR reduction method using null-constrained peak cancellation signal for massive MIMO-OFDM systems," *IEEE Open Journal of the Communications Society*, vol. 5, p. 3544–3555, 2024, doi: 10.1109/OJCOMS.2024.3406819.
- [52] J. Saito, T. Hara and K. Higuchi, "Peak cancellation signal-based parallel PAPR reduction method using low-dimensional null space in massive MIMO-OFDM," in *2024 International Conference on Computing, Networking and Communications (ICNC)*, 2024, pp. 361–365, doi: 10.1109/ICNC59896.2024.10556076.
- [53] J. Saito, N. Nonaka and K. Higuchi, "Peak cancellation signal generation considering variance in signal power among transmitter antennas in PAPR reduction method using null space in MIMO channel for MIMO-OFDM signals," *IEICE Transactions on Communications*, early access, doi: 10.23919/transcom.2023EBP3125.
- [54] A. Arvola, S. K. Joshi, A. Tölli and D. Gesbert, "PAPR reduction in MIMO-OFDM via power efficient transmit waveform shaping," *IEEE Access*, vol. 10, p. 47906–47920, 2022, doi: 10.1109/ACCESS.2022.3172325.
- [55] C. Ni, Y. Ma and T. Jiang, "A novel adaptive tone reservation scheme for PAPR reduction in large-scale multi-user MIMO-OFDM systems," *IEEE Wireless Communications Letters*, vol. 5, no. 5, p. 480–483, Oct. 2016, doi: 10.1109/LWC.2016.2588489.
- [56] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu and F. Rusek, "A low-complex peak-to-average power reduction scheme for OFDM based massive MIMO systems," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2014, pp. 114–117, doi: 10.1109/ISCCSP.2014.6877829.
- [57] S. Muneer, J. R. Sanchez, L. Van der Perre, O. Edfors, H. Sjöland and L. Liu, "System design and performance for antenna reservation in massive MIMO," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, 2022, pp. 1–5, doi: 10.1109/VTC2022-Fall57202.2022.10012708.
- [58] M. Q. Khan, M. D. Nisar and H. A. Khan, "On PAPR reduction and transmit distortion compensation in massive MIMO OFDM systems," in *2021 IEEE 93rd*

- Vehicle Technology Conference (VTC2021-Spring)*, 2021, pp. 1–5, doi: 10.1109/VTC2021-Spring51267.2021.9448801.
- [59] T. Kageyama, O. Muta and H. Gacanin, "Enhanced peak cancellation with simplified in-band distortion compensation for massive MIMO-OFDM," *IEEE Access*, vol. 8, p. 73420–73431, 2020, doi: 10.1109/ACCESS.2020.2986280.
- [60] R. Nojima, O. Muta, T. Makita and K. Maruta, "Joint peak cancellation and in-band distortion compensation scheme for precoded massive MIMO-OFDM with uniform planar array," in *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2022, pp. 301–306, doi: 10.1109/PIMRC54779.2022.9977810.
- [61] Ö. T. Demir, E. Björnson and L. Sanguinetti, "Foundations of user-centric cell-free massive MIMO," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, p. 162–472, 2021, doi:10.1561/2000000109.
- [62] Ł. Skomra, B. Jelonck and K. Staniec, "On the impact of residual transmitter distortions in 5G massive MIMO base stations," in *2023 International Wireless Communications and Mobile Computing (IWCMC)*, 2023, pp. 380–385, doi: 10.1109/IWCMC58020.2023.10182434.
- [63] 3GPP, "NR; Physical channels and modulation (Release 18)," TS 38.211, Jun. 2024. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.211/38211-i30.zip.
- [64] R. G. Lyons, *Understanding Digital Signal Processing*, 2nd ed., Prentice Hall PTR, 2004.
- [65] 3GPP, "NR; Base Station (BS) radio transmission and reception (Release 18)," TS 38.104, Jun. 2024. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.104/38104-i60.zip.
- [66] 3GPP, "NR; Base Station (BS) conformance testing; Part 2: Radiated conformance testing (Release 18)," TS 38.141-2, Jun. 2024. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.141-2/38141-2-i60.zip.
- [67] T. Araujo and R. Dinis, "On the accuracy of the gaussian approximation for the evaluation of nonlinear effects in OFDM signals," *IEEE Transactions on*

- Communications*, vol. 60, no. 2, p. 346–351, Feb. 2012, doi: 10.1109/TCOMM.2011.102011.110151.
- [68] ZTE, "Discussion on DL 1024QAM for NR FR1," 3GPP TDoc R1-2007977, 2020. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/TSG_RAN/WG1_RL1/TSGR1_103-e/Docs/R1-2007977.zip.
- [69] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (Release 18)," TR 38.901, Mar. 2024. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/38901-i00.zip.
- [70] 3GPP, "NR; Physical layer procedures for data (Release 18)," TS 38.214, Jun. 2024. Accessed: Sep. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.214/38214-i30.zip.
- [71] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, p. 2921–2935, May 2017, doi: 10.1109/TWC.2017.2672540.
- [72] I. Kim, J. Um and S. Park, "Implementation and performance evaluation of 256-QAM in vienna system level simulator," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*, 2018, pp. 556–559, doi: 10.23919/ICACT.2018.8323830.
- [73] Ł. Skomra, B. Jelonnek and K. Staniec, "Is antenna reservation superior to increasing input back-off in 5G massive MIMO base stations?," in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023, pp. 1–6, doi: 10.1109/VTC2023-Fall60731.2023.10333415.
- [74] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed., Wiley-Interscience, 2005.
- [75] G. F. Masters and S. F. Gregson, "Coordinate system plotting for antenna measurements," in *Antenna Measurement Techniques Association (AMTA) Annual Symposium*, 2007.

List of Figures

Figure 4.1: Base station, using an array of M antennas, transmitting distorted signals to K single-antenna users. _____	28
Figure 4.2: Normalized gain of the per antenna precoding weights for 3 UEs in multi-path LoS channels: (a) equal per UE power; (b) power allocated in a way to compensate the differences between path losses. _____	30
Figure 4.3: Power distribution between J subcarriers and K users: (a) example of an equal power distribution between users, equalizing channel gains; (b) equal power distribution between both users and subcarriers. _____	31
Figure 4.4: Block diagram of iterative clipping and filtering (ICF). _____	32
Figure 4.5: Clipping level (CL) and input back-off (IBO) in relation to average and peak output powers of a linearized power amplifier (PA) and the nominal peak-to-average power ratio (PAPR). _____	33
Figure 4.6: Tolerable EVM from distortion reduces when the headroom between the UE SINR and the required SNR decreases. _____	35
Figure 4.7: Normalized array factor (AF) for a uniform rectangular array (URA) with 32 rows and columns, steered to $\phi = 30^\circ$ and $\theta = -15^\circ$: (a) AF azimuth cut at $\theta = -15^\circ$; (b) AF, excluding values below -40 dB. _____	36
Figure 4.8: Radiation patterns of the wanted signal x and clipping distortion d for MU-MIMO transmission over free-space LoS channels to 2 UEs located at azimuth angles of: a) -30° and 10° ; b) -30° and 30° . _____	38
Figure 4.9: CCDF of generated NR20 data closely follows the theoretical Rayleigh distributed magnitudes. _____	40
Figure 4.10: The power of error resulting from peak limitation reduces with increasing clipping ratio (CR). _____	41
Figure 4.11: The theoretical power ratios of the user-directed and other spatially filtered third-order intermodulation products (IMPs) to all the third-order IMPs. _____	42
Figure 4.12: The uniformly radiated uncorrelated noise limits the reduction of EVM when increasing the number of UEs K , depending on the number of antennas M . _____	44
Figure 4.13: Expected radiation patterns of the wanted signal x as well as the spatially filtered and the explicitly user-directed third-order intermodulation products (IM3s), for two different sets Γ of phase gradients $\Delta\gamma$. _____	46
Figure 4.14: Normalized radiation patterns, excluding values below -10 dB, for the expected spatially filtered third-order intermodulation products of 16×16 URA for 4 different sets of phase gradients (with beams pointing directions for the wanted data signal visualized with crosses). _____	49

- Figure 4.15: Comparison of normalized radiation patterns, excluding values below -10 dB, of spatially filtered third-order intermodulation products radiated from 32×32 and 8×8 rectangular arrays, for 2 sets of steering angles for 4 UEs (with directions of UEs marked with crosses). 51
- Figure 4.16: Radiation patterns of the wanted signal x and distortion d for uniform linear array (ULA) with 64 antennas, using directional antenna elements with 8 dBi, for steering to 30° . 54
- Figure 4.17: Prediction of per antenna EVM compared with the hard clipping (HC) and the iterative clipping and filtering (ICF) simulation results. 59
- Figure 4.18: EVM statistics (RMS: solid; RMS \pm standard deviation: filled ranges) for the best- and worst-cases of UE locations in free-space LoS channels. 62
- Figure 4.19: Comparison of EVM simulation results for multi-path with LoS (RMS: solid; RMS \pm std dev: filled ranges) and free-space LoS, for UEs located at the sets of angles of the best- and worst-case scenarios. 64
- Figure 4.20: Prediction of the portion of spatially filtered intermodulation products (IMPs) received by UEs, based on numerical simulations of $1e6$ Monte Carlo trials. 66
- Figure 4.21: Expected EVM distribution, over spatial scenarios, for a free-space LoS channel for different number of UEs K for 32 (solid lines) and 128 (dashed lines) antennas M . Based on the numerical simulations of third-order IMPs, assuming -23.5 dB EVM for $K = 1$. 67
- Figure 4.22: Comparison for the statistics of EVM (RMS of per UE EVMs), over spatial scenarios, of link-level simulation results (solid lines) with the prediction based on the numerical simulations using precoding weights (dash-dotted lines), for a free-space LoS channel for different number of UEs K for 32 antennas. 68
- Figure 4.23: The differences of percentiles of EVM between the link-level simulations and precoding-based prediction for different number of UEs (K) for transmission using 32 antennas in deterministic LoS and frequency-selective channels with LoS. 69
- Figure 4.24: Comparison of the statistics over spatial scenarios of RMS of per UE EVMs, for link-level simulation results (solid lines) with the numerical simulations for the prediction based on precoding (dash-dotted lines), for a free-space LoS channel for different number of UEs K for a square URA with 64 antennas. 71
- Figure 4.25: In the worst case of equal elevation steering (here $\theta = 0^\circ$) for all UEs, the EVM statistics for URA depend only on the number of columns M_y , not on the total number of antennas M (here 256 for URA). Based on the EVM prediction from the numerical simulations of IMPs, assuming -23.5 dB EVM for $K = 1$. 71
- Figure 4.26: Comparison of the expected EVM distribution for URAs with 16 and 32 columns M_y (both with 8 rows). Based on the numerical simulations of predicted 3rd-order IMPs, assuming -23.5 dB EVM for $K = 1$. 72

- Figure 4.27: Statistics of EVM (RMS of per UE EVMs, covering transmission of 10 slots) over 100 spatial scenarios for link-level simulation results for CDL-A (NLoS) channels for different number of UEs K for 32 antennas. _____ 73
- Figure 4.28: Comparison of the radiation patterns of the wanted signal x for equal power and equal SNR allocation strategies, for linear array with 32 antennas. Uniformly radiating distortion d shown for reference. _____ 75
- Figure 4.29: Comparison of the radiation patterns of the wanted signal x and nonlinear distortion d for equal power and equal SNR allocation strategies, for linear array with 32 antennas. _____ 77
- Figure 5.1: The increase in the spectral efficiency (SE) per resource element (RE) with increasing signal-to-interference-plus-noise-plus-distortion ratio (SINDR), bounded by the channel capacity, is limited in practice by the maximum supported modulation order and code rate. _____ 83
- Figure 5.2: Maximum cell throughput for 100 MHz 5G NR carrier for a range of antennas M in relation to the number of UEs K , each with SINR of 1 when $M = K$. Assuming 80% resources available for downlink data transmission and spectral efficiency limited by 1024QAM at 948/1024 code rate. _____ 84
- Figure 5.3: Examples of frequency distributions of SNR for an equal level of 20 dB wideband SNR. _____ 85
- Figure 5.4: Example of the mean channel gain and gains of the channel responses of 3 antennas out of 64. _____ 86
- Figure 5.5: Channel hardening measure, from (5.3), per subcarrier for simulated spatially correlated NLoS channels improves for higher number of antennas M . Based on a figure originally published in [62] (co-authored by the PhD candidate). _____ 87
- Figure 5.6: Minimum ratios between the wanted signal power (S for received, S_m for per antenna) and the power of the additive Gaussian noise (representing either the noise power at the receiver N or the per TX antenna distortion power D_m) required for 10% BLER in CDL-C (NLoS) channel depend on the number of antennas M . Based on a figure originally published in [62] (co-authored by the PhD candidate). _____ 90
- Figure 5.7: Comparison of the sum user data rates for 4 UEs located at different distances from the base station (50, 75, 100, and 500 m) for 2 different power allocation strategies, with no distortion and with iterative clipping and filtering (ICF) distortion from Table 4.8, for linear array with different number of M antennas. _____ 92
- Figure 5.8: Per user data rates (each color represents specific UE) for two different power allocation strategies, with no distortion and with iterative clipping and filtering (ICF) distortion, as a function of the SNR of the UE closest to the base station when equal power allocation is used. _____ 93
- Figure 5.9: Cumulative distribution functions (CDFs) of signal-to-interference-plus-noise ratio (SINR), with multicell minimum mean-squared error (M-MMSE) and regularized zero forcing

(RZF) precoding, for 16-cell network with 10 UEs and base station with M antennas in each cell.

- _____ 94
- Figure 5.10: Sum data rate per resource element (RE) for M-MMSE precoding, depending on the number of antennas, averaged over cells for a range of per antenna EVM levels. _____ 95
- Figure 5.11: Larger input back-off (IBO) leads to higher power usage, but beyond some level it may result in diminishing returns from the perspective of the cell throughput performance. _____ 98
- Figure 5.12: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses. _____ 101
- Figure 5.13: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses. _____ 102
- Figure 5.14: Normalized radiation patterns of spatially filtered third-order intermodulation products (IMPs), with each color level representing a 1 dB range between -10 and 0 dB. Directions of UEs marked with crosses. _____ 104
- Figure 5.15: Comparison of the achievable per resource element (RE) sum data rates with the number of UEs going from 3 to 8, for distortion-unaware scheduling from Figure 5.13 and distortion-aware from Figure 5.14. _____ 106
- Figure 6.1: Comparison of the complementary cumulative distribution functions (CCDFs) of the instantaneous power of the data and compensation signals for either 16 or 8 out of 64 antennas reserved, at the input and output of the iterative clipping and filtering (ICF) block. Based on a figure originally published in [73] (co-authored by the PhD candidate). _____ 112
- Figure 6.2: Comparison of radiation patterns for a linear array of 32 omni-directional antennas, with a beam-steering to 30° azimuth angle for no tapering, as well as tapering based on Hamming and Hann windows. _____ 113
- Figure 6.3: A comparison of the gains of tapering weights based on Hamming and Hann windows: (a) per antenna gain for a linear array of 32 antennas; (b) cumulative distribution functions (CDFs) of the per antenna gains, depending on the number of antennas M in a linear array. _____ 114
- Figure 6.4: Block diagram of the proposed compensation approach using antennas with a power headroom. Only selected elements of the processing chain, relevant from the perspective of the proposed method, shown. _____ 116
- Figure 6.5: Block diagram of the compensation calculation block, from Figure 6.4, for a single subcarrier. _____ 118
- Figure 6.6: The peak power capability required for an array of 32 antennas using Hamming-based tapering to achieve the same error level and mean power at the user as for a non-tapered array reduces with increasing the number of compensating antennas MC . _____ 120

- Figure 6.7: Boxplots representing the statistics of EVM over UEs, for the scenarios from Table 6.1, achieved with a non-tapered array as well as a tapered array without compensation and with compensation, using zero forcing (ZF) and maximum ratio (MR) precoding of the compensation signals. _____ 123
- Figure 6.8: Comparison of the per antenna distribution of the peak and average power levels at the input and output of the iterative clipping and filtering (ICF) block for tapered antenna array without and with compensation of the PAPR reduction distortion (in 50% of the antennas, located symmetrically at the edges). _____ 125
- Figure 6.9: Comparison of the performance for tapered antenna array, in the 6 UE CDL-D (LoS) channel case from Table 6.2, for compensated and uncompensated array (with either equal peak or PAPR limits, over antennas): (a) energy efficiency; (b) power usage. _____ 126
- Figure A.1: Horizontally oriented uniform linear array (ULA), steered to azimuth angle $\phi = 45^\circ$. _____ 148
- Figure A.2: Comparison of radiation patterns for uniform linear arrays (ULAs) with 8 and 32 antennas, steered to azimuth angle ϕ of 45° . _____ 149
- Figure A.3: Comparison of normalized array factors (AFs) for uniform linear arrays (ULAs) with different spacings d between 32 antennas, steered to $\phi = -60^\circ$. _____ 151
- Figure A.4: The relation of the azimuth steering angle and the related phase gradient, for $d = \lambda/2$. _____ 152
- Figure A.5: Normalized array factor (AF) for uniform linear array (ULA) with $d = \lambda/2$ spacing between 32 antennas, with the phase gradient sweeping from 0° to 360° , with a 60° step. Solid line represents the part of AF in the $\pm 90^\circ$ field of view of the array, while dotted line shows the part of AF beyond that field of view. _____ 152
- Figure A.6: Uniform rectangular array (URA) placed in the y - z plane, steered to azimuth ϕ and elevation θ . _____ 153
- Figure A.7: Radiation patterns for arrays with half-wavelength spacing, steered to azimuth angle $\phi = 45^\circ$ and elevation angle $\theta = 20^\circ$: URA with isotropic sources in (a) and (b); ULA with directional antennas (8 dBi) in (c) and (d). In (b) and (d), power levels 40 dB lower than maximum ignored. _____ 155
- Figure B.1: Power spectrum for the passband representation of the transmitted signal, with 1 to 3 input tones passed through a third-order nonlinearity. Input signal x and error e (difference between equalized output and the input) is visualized. _____ 162
- Figure B.2: Power spectrum for the passband representation of the transmitted signal, for the input signal x and error signal e (difference between equalized output and the input) for: (a) x with 90, 110 and 120 MHz tones; (b) x with 90, 110 and 130 MHz tones; (c) a multi-tone signal, covering 40 MHz, with 1 MHz equidistant spacing between tones. _____ 164

- Figure B.3: Radiation patterns of wanted signal x , as well as user-directed and spatially filtered third-order IMPs (IM3s): (a) 32 antennas; (b) 128 antennas. _____ 166
- Figure B.4: Radiation patterns of wanted signal x , as well as user-directed and spatially filtered third-order IMPs (IM3s). _____ 167
- Figure B.5: Radiation patterns, for linear array with 32 antennas, for all the identified unfavorable locations of adding the 3rd beam. _____ 168
- Figure C.1: Comparison of BLER vs wideband signal-to-noise-plus-distortion ratio (SNDR) for SISO transmissions in CDL-C (NLoS) and CDL-D (LoS) channels. Results for Gaussian noise added either on TX or RX side. Based on a figure originally published in [62] (co-authored by the PhD candidate). _____ 171
- Figure D.1: A comparison of the radiation patterns of the wanted signal x and distortion d , based on link-level simulation results, with the expectations of the third- and fifth-order intermodulation (IM) beams calculated based on precoding weights. UEs are located at the azimuth angles marked by the vertical dashed lines. _____ 173
- Figure D.2: Comparison of EVM simulation results for NLoS channels: CDL-A (RMS: solid; RMS \pm std dev: filled ranges), as well as CDL-A and CDL-C with clipping ratios (CRs) scaled with per antenna norms of precoding weights, for UEs located at the sets of angles of the best- and worst-case scenarios for LoS. _____ 174
- Figure D.3: For 1 UE, when the number of antennas M is increased (from 1 to 512), the EVM resulting from hard clipping (HC) changes differently depending on the channel type. Based on a figure originally published in [62] (co-authored by the PhD candidate). _____ 176
- Figure D.4: Examples of the EVM vs the number of UEs K for the quantization noise model. Lines represent the mean EVM over all UEs, while the filled ranges show the EVM variation between UEs. Based on a figure originally published in [62] (co-authored by the PhD candidate). __ 178

List of Tables

Table 4.1: The ratio U , of the user-directed IMPs to all the 3rd-order IMPs, from (4.16). _____	38
Table 4.2: The relation between the linearly scaled clipping ratio (CR) and the EVM prediction. _____	41
Table 4.3: The relation between the clipping ratio (CR) and the error vector magnitude (EVM). _____	59
Table 4.4: Definition of the best- and worst-case scenario examples for ULA with 32 antennas. _____	61
Table 4.5: RMS of EVM (RMS over UEs per slot) over 100 slots for free-space LoS. _____	62
Table 4.6: Predicted improvement of the per user EVM in relation to the per antenna EVM. _____	67
Table 4.7: Examples of per UE transmit powers and EVMs, for two different power allocation strategies: equal power; power distribution based on path loss, to achieve equal SNR. _____	76
Table 4.8: Per user EVMs resulting from distortion introduced by ICF, for UEs in free-space LoS channels, depending on the power allocation strategy and the number antennas M . _____	77
Table 5.1: The relation between the input back-off (IBO) and the per antenna EVM, the resulting signal-to-distortion ratio (SDR) for 64 antennas M and 10 users K , and the sum data rate per resource element (RE) averaged over cells for M-MMSE, RZF, and MR precoding. _____	97
Table 5.2: Ranges between percentiles of the EVM performance statistics over sets of UE locations, based on the free-space LoS prediction results from Figure 4.21. _____	100
Table 5.3: Summary of effective array gains and statistics of signal-to-distortion ratios (SDRs) for the UEs located at the directions from Figure 5.12 and Figure 5.13. _____	103
Table 5.4: Summary of effective array gains and statistics of signal-to-distortion ratios (SDRs) for the UEs located and the directions from Figure 5.14. _____	105
Table 6.1: Comparison of EVMs with and without compensation of transmitter distortion. _____	121
Table 6.2: EVMs for a tapered array with 0 or 16 antennas out of 32 in the compensating set \mathcal{C} . _____	123
Table B.1: Intermodulation products (IMP) from (B.4), grouped in $2A-B$ and $A+B-C$ classes. _____	161
Table B.2: List of the frequencies of the IMP tones for the 3-tone input from Figure B.1(c). _____	163
Table B.3: List of IMPs for a 3-tone input with equidistant frequency spacing. Rows sorted with increasing frequency. _____	163
Table B.4: Summary of phase gradients and resulting steering angles of the spatially filtered intermodulation products (IMPs) for UEs placed at azimuth angles of -40° , 5° , and 30° . _____	166
Table B.5: Phase gradients and steering angles for identified intermodulation beams. _____	168
Table C.1: SNDR for 10% BLER for SISO transmission using 256QAM and 64QAM at different code rates. _____	171
Table D.1: RMS over 100 slots of RMS EVMs (over UEs per slot) for CDL-A (NLoS). _____	173
Table D.2: Improvement in EVM caused by adjustment of the per antenna CRs. _____	175
Table D.3: Overview of the EVM statistics of median EVMs over UEs, based on results in [62]. _____	177

Appendix A

Antenna Arrays

The geometry and orientation of the antenna arrays, as well as their relation with azimuth and elevation angles, are described here as used in this dissertation, to avoid ambiguity. The description of the beam steering forms a basis for understanding about the way the nonlinear distortion is expected to radiate from the array, which is covered in Appendix B.

In this dissertation, linear and rectangular (i.e., planar) antenna arrays are considered. From the hardware design perspective, the arrays utilized here follow the definition of the term *uniform array* from [74, Sec. 6.3], by consisting of uniformly spaced identical antenna elements, with identical magnitude and progressive phase applied to the antenna array input signals, in the analog part of the TX processing chain. However, the selected phase progression is 0° and both magnitudes and phases of the signals at the inputs to antennas are controlled by precoding weights applied in the digital signal processing part of the TX chain, individually to each of the subcarriers of the OFDM symbol.

While the antenna arrays utilized in Massive MIMO base stations radiate the signals in alignment with the estimated UE channel responses, using linear precoding, which may lead to radiation patterns that are far from single beams pointing in particular directions, especially for UEs in NLoS channels, it is vital to consider the principles of radiation from antenna arrays using beam steering. After introducing the ULA with horizontal steering in Section A.1, the general case of URA is described in Section A.2. Finally, beamforming using precoding weights (with the amplitude controlling the shape of the radiation pattern and the phase impacting the scanning capabilities [74, Sec. 16.7.2]) is covered in Section A.3.

Note that all the calculations in this appendix relate with the geometry, orientation in space, and definition of angles that is aligned with the rest of this dissertation. For a different orientation of the array in space or relation with angles, the provided equations may need to be adjusted, for instance based on [74].

A.1. Uniform Linear Array with Horizontal Steering

A.1.1. Geometry, Orientation and Beam Steering

A horizontally oriented ULA, placed along the y -axis, allows to steer the beam to different azimuth angles ϕ , but has no elevation angle θ steering capability. The array consists of M antennas, uniformly spaced at distance d from each other, as visualized in Figure A.1.

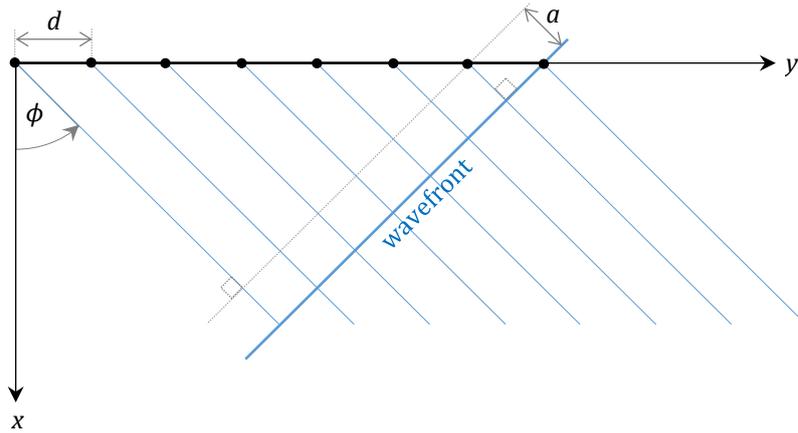


Figure A.1: Horizontally oriented uniform linear array (ULA), steered to azimuth angle $\phi = 45^\circ$.

For the UE located at azimuth angle ϕ and elevation angle $\theta = 0^\circ$, the phase shifts of antennas must be selected such that the wavefront propagates in the UE direction. For ULA, this is achieved by applying a linear phase gradient (i.e., a progressive phase shift) of $\Delta\gamma$ so that the signal is appropriately delayed in different antennas.

In case of a positive azimuth steering angle ϕ , the leftmost antenna in Figure A.1 must radiate first. The signal in all the other antennas must be delayed, using a progressive phase $\Delta\gamma$, so that the rightmost antenna radiates as the last one, in this case. To steer the beam to negative azimuth angles the delays of the antennas must be reversed.

The required phase gradient $\Delta\gamma$ can be calculated based on the time delay required between antennas for coherent combination in the direction of the steering angle ϕ . The difference in the distance a that need to be traveled by the signals radiated from adjacent antennas, so that they form a wavefront in the required direction, as visualized with an example in Figure A.1, is

$$a = d \sin \phi. \quad (\text{A.1})$$

Hence the time difference between signals on adjacent antennas should be

$$\Delta t = \frac{a}{c} = \frac{d \sin \phi}{c} \quad (\text{A.2})$$

where c is the speed of light.

When starting from the antenna closest to the origin in Figure A.1, the progressive phase shift between antennas, required to approximate the desired progressive time delay of $-\Delta t$ is

$$\Delta\gamma = -2\pi f\Delta t. \quad (\text{A.3})$$

By plugging (A.2) into (A.3) and defining the frequency in terms of the wavelength (i.e., $f = c/\lambda$) the phase gradient $\Delta\gamma$ can be described in terms of the distance between antennas d and the steering angle ϕ as

$$\Delta\gamma = -\frac{2\pi}{\lambda}d \sin \phi. \quad (\text{A.4})$$

Let κ be the angular wavenumber associated with the linear wavelength λ , such that

$$\kappa = \frac{2\pi}{\lambda}. \quad (\text{A.5})$$

The phase gradient can be described in terms of the wavenumber as

$$\Delta\gamma = -\kappa d \sin \phi. \quad (\text{A.6})$$

For example, the linear phase gradient $\Delta\gamma$ required to steer a ULA with a half-wavelength spacing between antennas to azimuth angle ϕ of 45° is about -127° , irrespective of the number of antennas M in the array, as it follows from (A.6). Radiation patterns for ULAs with 8 and 32 omni-directional antennas for that linear phase gradient are demonstrated in Figure A.2, for the same sum output power over antennas of 30 dBm for both arrays. It can be noticed that an increase of the number of antennas leads to a reduction of width of both the main lobe and the side lobes. It can be additionally observed that the null angles of the array with 8 antennas are a subset of the null angles for the array with 32 antennas.

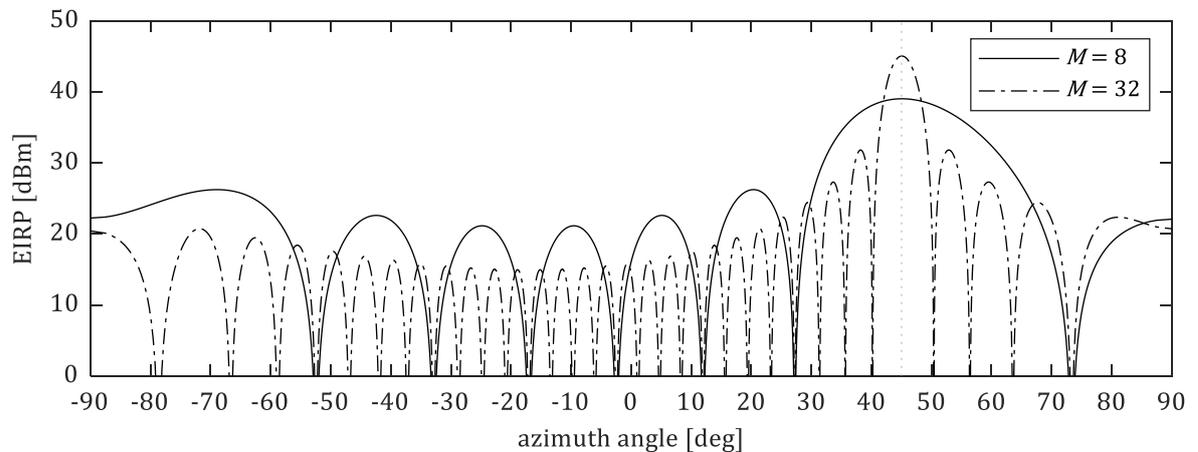


Figure A.2: Comparison of radiation patterns for uniform linear arrays (ULAs) with 8 and 32 antennas, steered to azimuth angle ϕ of 45° .

A.1.2. Array Factor

The array factor (AF) describes the radiation pattern of an array of isotropic sources. The AF for horizontally-oriented ULA for the azimuth cut at $\theta = 0^\circ$ and a uniform amplitude excitation can be expressed, adapting the equation from [74, Sec. 6.6], as

$$\text{AF}(\psi) = \sum_{m=1}^M e^{j(m-1)\psi} \quad (\text{A.7})$$

where M is the number of antennas and

$$\psi = \kappa d \sin \phi + \Delta\gamma. \quad (\text{A.8})$$

The angular directions of the maximum radiation power can be calculated based on (A.8), by setting it to 0, leading to fully coherent combining with a magnitude of $|\text{AF}(0)| = M$. Since the array factor in (A.7) is a periodic function of ψ [74, Sec. 7.4.2], repeating every 2π radians, it follows that the AF is maximized for phase gradients of the following form:

$$\Delta\gamma = -\kappa d \sin \phi + n2\pi \quad (\text{A.9})$$

with $n = (0, \pm 1, \pm 2, \dots)$, where $n = 0$ defines the direction of the principal maximum.

Without any beam steering applied (i.e., $\Delta\gamma = 0$), the main beam is visible at an observation angle $\phi = 0^\circ$, which is the mechanical boresight direction (along the x -axis here) of the array. In that case, the visibility of any additional maxima of AF (i.e., the directions that the grating lobes are pointing), resulting from phase gradients of $\Delta\gamma = n2\pi$ for $n \neq 0$, depends on the spacing between antennas.

For $d < \lambda/2$, there is only a single maximum of AF visible in the range of azimuth observation angles between -90° and 90° , irrespective of the steering angle. However, for $\lambda/2 \leq d < \lambda$, additional maxima of AF (grating lobes) can be experienced in that range, depending on both the spacing d and the steering angle [74, Sec. 6.10.1].

Assuming the azimuth steering for the Massive MIMO base station to be in practice limited to $\pm 60^\circ$, the related range of the potentially applied phase gradients is further limited for $d = \lambda/2$ from $\pm 180^\circ$ down to about $\pm 156^\circ$. However, when the distance between antennas is increased, a grating lobe may appear within the steering range of angles. For example, when the main beam is steered to $\phi = -60^\circ$ a grating lobe (with the same power) will appear at the other end of the steering range (i.e., at $\phi = 60^\circ$) for antenna spacing d for which the phase gradient is $-\pi$, which is

$$d = \frac{-\Delta\gamma\lambda}{2\pi \sin \phi} \Big|_{\Delta\gamma=-\pi} = \frac{\lambda}{2 \sin \phi} \Big|_{\phi=60^\circ} \approx \frac{\lambda}{1.73}. \quad (\text{A.10})$$

This is demonstrated with a normalized AF (i.e., AF divided by the number of antennas M) for ULA with 32 antennas steered to $\phi = -60^\circ$, visualized in Figure A.3, where for d of about $\lambda/1.73$ a grating lobe at an azimuth angle of 60° is visible. While for $d = \lambda$, a grating lobe at about 7.7° can be observed.

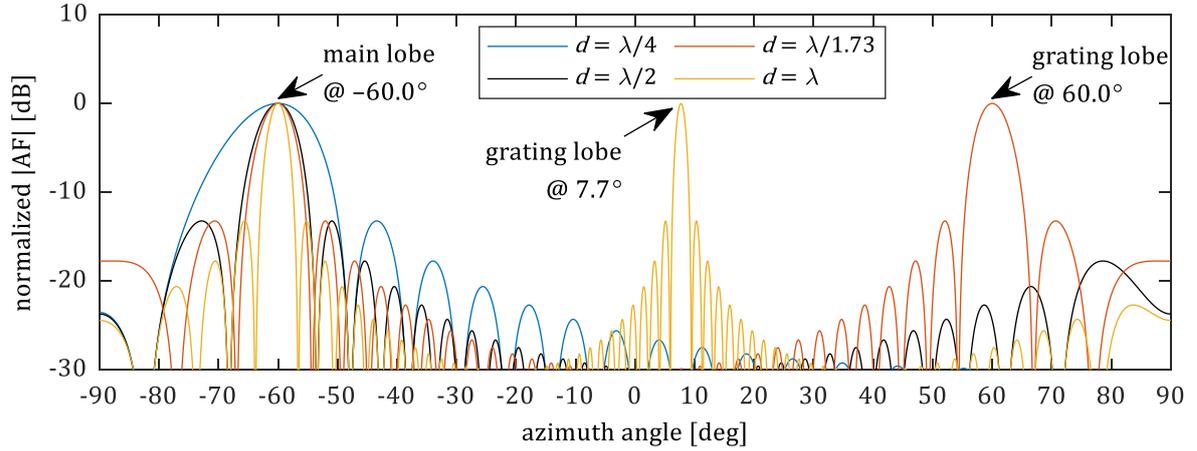


Figure A.3: Comparison of normalized array factors (AFs) for uniform linear arrays (ULAs) with different spacings d between 32 antennas, steered to $\phi = -60^\circ$.

A.1.3. Critically Spaced Array

For an array with a critical spacing between antennas of half the wavelength [1, Sec. 7.4.2], the calculation in (A.9) for the phase gradient $\Delta\gamma$ required to steer the beam to ϕ simplifies to

$$\Delta\gamma = -\pi \sin \phi + n2\pi. \quad (\text{A.11})$$

It follows from (A.11) that a grating lobe for $d = \lambda/2$ would appear in the resolvable range of angles, only if an extreme steering angle of $\pm\pi/2$ radians was used. This is because for $|\phi| < \pi/2$ the phase gradient for the principal maximum of the AF (i.e., for $n = 0$) is $|\Delta\gamma| < \pi$, hence for any $|n| > 0$ the resulting phase gradient $|\Delta\gamma| > \pi$, which leads to an angle that is beyond the visible region of the array with $d = \lambda/2$.

In order to identify the steering angle ϕ based on the applied linear phase gradient $\Delta\gamma$ the equation in (A.11) can be rewritten as follows:

$$\phi = \arcsin\left(-\frac{\Delta\gamma + n2\pi}{\pi}\right) = \arcsin\left(-\frac{\Delta\gamma}{\pi} - 2n\right). \quad (\text{A.12})$$

The relationship between the phase gradient $\Delta\gamma$ and the steering angle ϕ for $n = 0$, visualized in Figure A.4, reveals that the range of phase gradient of $[-\pi, \pi]$ radians covers the azimuth steering range of $[-\pi/2, \pi/2]$ radians. In classical 3-sector cellular base

stations, the utilized azimuth steering range is considerably smaller, since it needs to cover a 120° -wide sector (leading to about $\pm 156^\circ$ range of practically usable phase gradients).

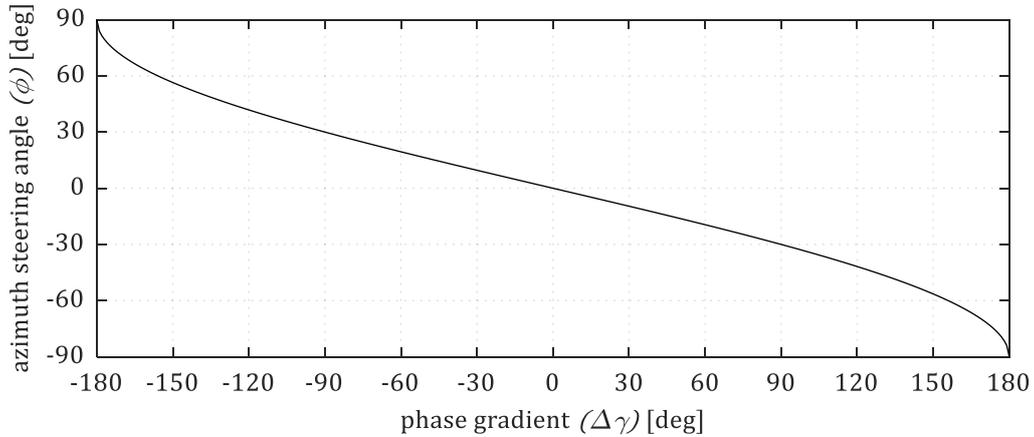


Figure A.4: The relation of the azimuth steering angle and the related phase gradient, for $d = \lambda/2$.

One can see what happens when the AF, from (A.7), is calculated for $d = \lambda/2$ over a full range of azimuth angles from $-\pi$ to π , as visualized in the examples in Figure A.5 where the phase gradient is swept from 0° to 360° with a step of 60° . Since the field of view of the array is limited to $\pm\pi/2$, the spatial images visible in AF at angles beyond that field of view are not radiated in practice.

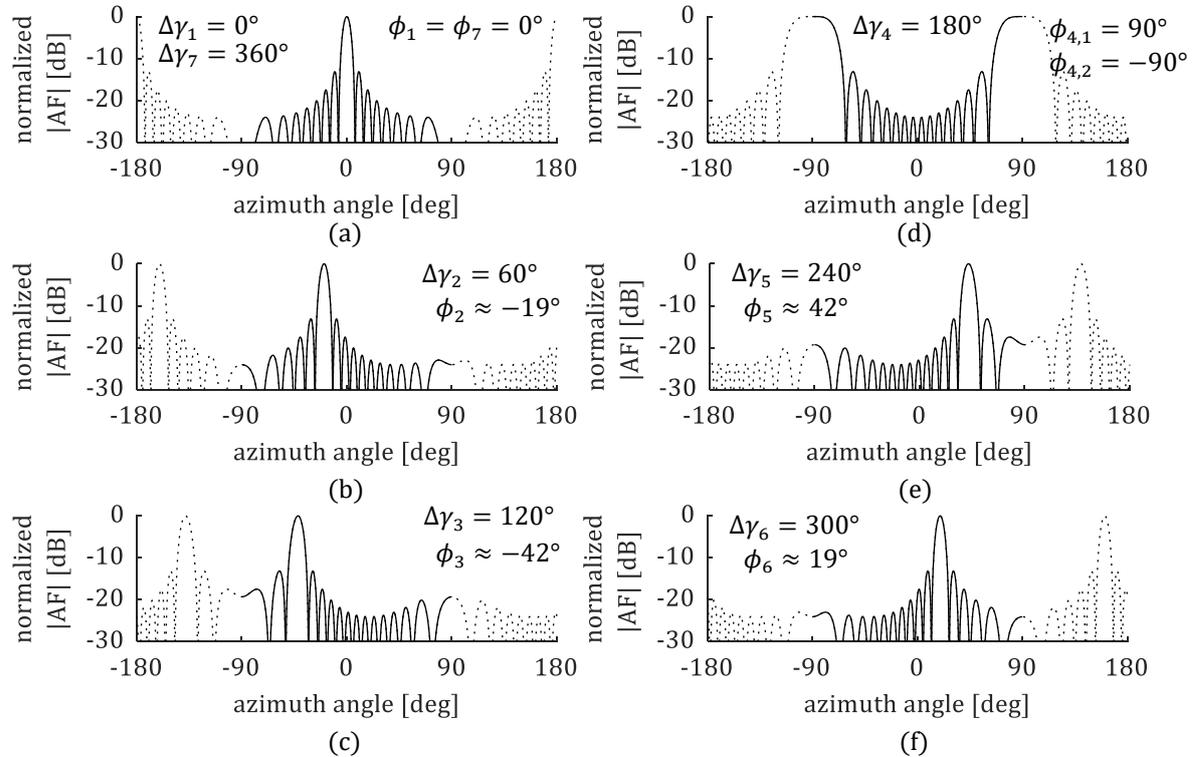


Figure A.5: Normalized array factor (AF) for uniform linear array (ULA) with $d = \lambda/2$ spacing between 32 antennas, with the phase gradient sweeping from 0° to 360° , with a 60° step. Solid line represents the part of AF in the $\pm 90^\circ$ field of view of the array, while dotted line shows the part of AF beyond that field of view.

Note that in order to radiate towards UE that is located at a non-zero elevation angle θ , the linear phase gradient $\Delta\gamma$ deviates from the value used to steer the beam to the azimuth angle ϕ of the UE. This is explained in the Section A.2, together with the extension to URAs.

A.2. Uniform Rectangular Array

A.2.1. Geometry, Orientation and Beam Steering

An URA, placed in the y - z plane, enables steering the beam to different azimuth ϕ and elevation θ angles. The array consists of M identical antennas, placed in M_z rows and M_y columns. The vertical spacing between the antennas (i.e., the distance between rows) is d_z , while the horizontal spacing between antennas (i.e., the distance between columns) is d_y . The positive azimuth angles go from the x -axis in the direction of the y -axis, while the positive elevation angles go from the x - y plane in the direction of the z -axis, as visualized in Figure A.6.

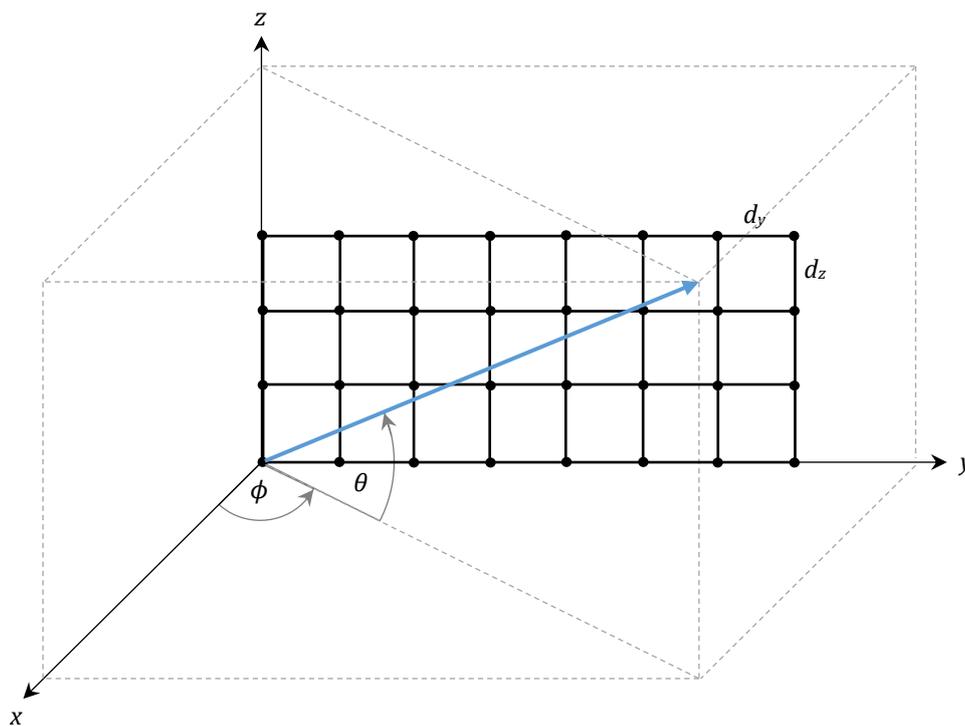


Figure A.6: Uniform rectangular array (URA) placed in the y - z plane, steered to azimuth ϕ and elevation θ .

The conversion from spherical coordinates to cartesian, for the orientation of axes and definition of angles from Figure A.6, can be done using the following transformation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos \phi \cos \theta \\ \sin \phi \cos \theta \\ \sin \theta \end{bmatrix}. \quad (\text{A.13})$$

The linear phase gradients are denoted as $\Delta\gamma_y$ for the columns of the array, distributed over the horizontal (y -axis) direction, and $\Delta\gamma_z$ for the rows of the array, distributed over the vertical (z -axis) direction.

The differences in the distance a that the signal needs to travel from adjacent antennas during a single period can be defined in relation to the spacing between antennas as follows:

$$\frac{a}{d_z} = \sin \theta \quad (\text{A.14})$$

$$\frac{a_{xy}}{d_y} = \sin \phi \quad (\text{A.15})$$

where a_{xy} is the distance travelled parallel to the x - y plane. The relation between a and a_{xy} is given by the elevation angle θ as follows:

$$\frac{a}{a_{xy}} = \cos \theta \quad (\text{A.16})$$

which leads to

$$\frac{a}{d_y} = \sin \phi \cos \theta. \quad (\text{A.17})$$

The related time differences in the propagation are as follows

$$\frac{c\Delta t}{d_y} = \sin \phi \cos \theta \quad (\text{A.18})$$

$$\frac{c\Delta t}{d_z} = \sin \theta. \quad (\text{A.19})$$

Similarly as for ULA, when starting from the antenna closest to the origin in Figure A.6, the approximation of the desired progressive time delay of $-\Delta t$ is done with progressive phase shifts between antennas, in both horizontal and vertical directions as

$$\Delta\gamma_y = -2\pi f \Delta t \quad (\text{A.20})$$

$$\Delta\gamma_z = -2\pi f \Delta t. \quad (\text{A.21})$$

By plugging Δt from (A.18) into (A.20) and from (A.19) into (A.21), defining the frequency in terms of the wavelength (i.e., $f = c/\lambda$) the horizontal and vertical phase gradients can be described in terms of the distances between antennas (d_y and d_z) and the steering angles (i.e., azimuth angle of ϕ and elevation angle of θ) as

$$\Delta\gamma_y = -\frac{2\pi}{\lambda}d_y \sin\phi \cos\theta = -\kappa d_y \sin\phi \cos\theta \quad (\text{A.22})$$

$$\Delta\gamma_z = -\frac{2\pi}{\lambda}d_z \sin\theta = -\kappa d_z \sin\theta. \quad (\text{A.23})$$

It can be observed that the phase gradient calculation for a horizontally-oriented ULA with azimuth steering, introduced in (A.4), is a special case of $\Delta\gamma_y$ when the elevation angle points to the horizon, with $\theta = 0$. With this kind of orientation of the array and definition of the azimuth and elevation angles, the horizontal phase gradient depends on both azimuth and elevation angles, while the vertical phase gradient $\Delta\gamma_z$ depends only on the elevation angle.

A comparison of the radiation patterns of URA (with $d_y = d_z = \lambda/2$) and ULA, steered to $\phi = 45^\circ$ and $\theta = 20^\circ$ is visualized in Figure A.7.

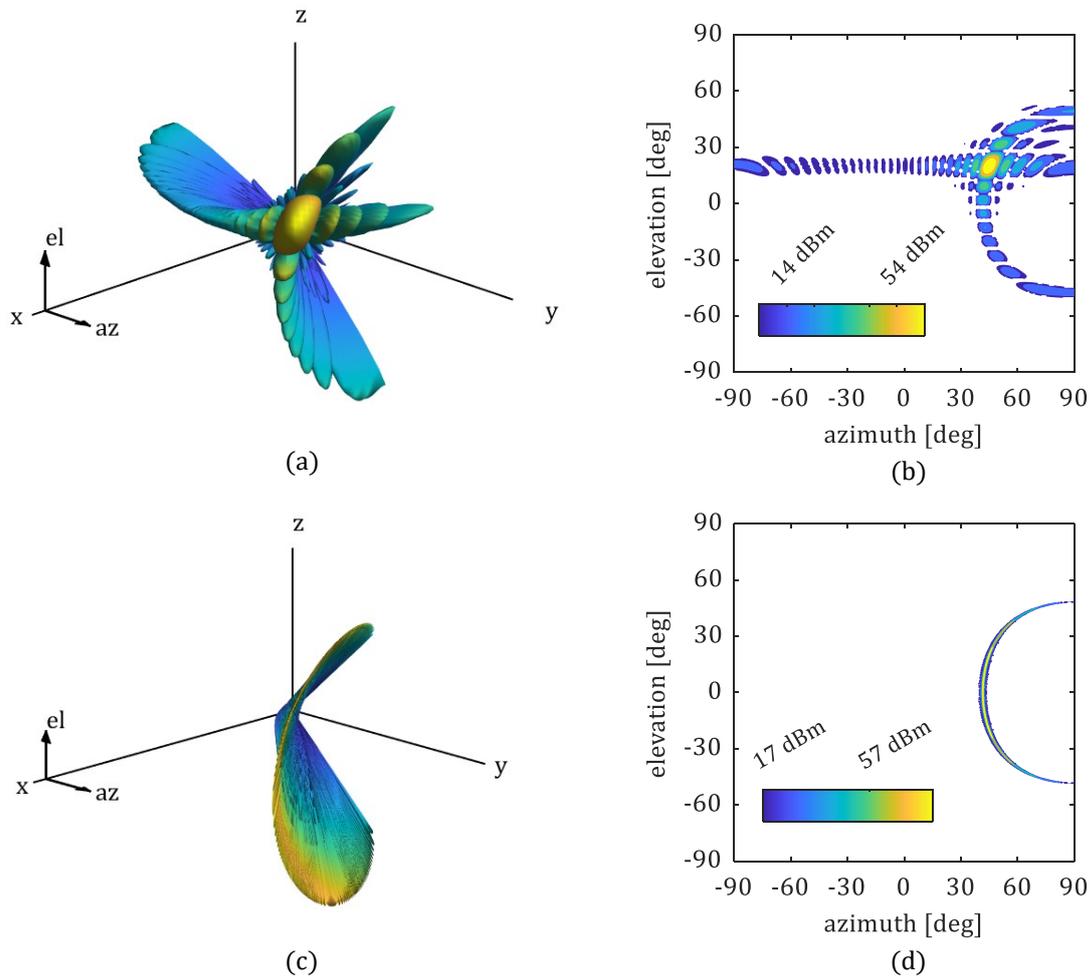


Figure A.7: Radiation patterns for arrays with half-wavelength spacing, steered to azimuth angle $\phi = 45^\circ$ and elevation angle $\theta = 20^\circ$: URA with isotropic sources in (a) and (b); ULA with directional antennas (8 dBi) in (c) and (d). In (b) and (d), power levels 40 dB lower than maximum ignored.

Steering an array with to $\phi = 45^\circ$ and $\theta = 20^\circ$, requires $\Delta\gamma_y$ of about -120° (which points to about 42° azimuth at 0° elevation), while $\Delta\gamma_z$ is about -62° deg. For an array of isotropic sources with $M_y = 32$, $M_z = 16$ this leads to the radiation pattern as visualized in Figure A.7(a). Alternatively, ULA with the same number of antennas (i.e., 512) and the same horizontal spacing, leading to the same $\Delta\gamma_y$ of about -120° using directional antennas (8 dBi in the boresight) results in the radiation pattern visualized in Figure A.7(c).

When comparing the 3D radiation pattern in Figure A.7(a) with the 2D radiation pattern in Figure A.7(b), the limitation of the latter one becomes apparent. Mapping a sphere to a surface may lead to not immediately intuitive result, depending on the projection method that is used [75].

A.2.2. Array Factor

The AF for URA with a uniform amplitude excitation can be expressed, adapting the equation from [74, Sec. 6.6], as

$$\text{AF}(\psi_y, \psi_z) = \sum_{m_y=1}^{M_y} e^{j(m_y-1)\psi_y} \sum_{m_z=1}^{M_z} e^{j(m_z-1)\psi_z} \quad (\text{A.24})$$

where M_y is the number of antennas in the horizontal direction (along the y -axis) and M_z is the number of antennas in the vertical direction (along the z -axis) and

$$\psi_y = \kappa d_y \sin \phi \cos \theta + \Delta\gamma_y \quad (\text{A.25})$$

$$\psi_z = \kappa d_z \sin \theta + \Delta\gamma_z. \quad (\text{A.26})$$

Similarly as for ULA, the angular directions of the maximum radiation power can be calculated based on (A.25) and (A.26), by setting both to 0, leading to fully coherent combining with a magnitude of $|\text{AF}(0,0)| = M_y M_z = M$. The AF is maximized for phase gradients of the following form:

$$\Delta\gamma_y = -\kappa d_y \sin \phi \cos \theta + n_y 2\pi \quad (\text{A.27})$$

$$\Delta\gamma_z = -\kappa d_z \sin \theta + n_z 2\pi \quad (\text{A.28})$$

with $n_y = (0, \pm 1, \pm 2, \dots)$ and $n_z = (0, \pm 1, \pm 2, \dots)$, where $n_y = n_z = 0$ defines the direction of the principal maximum.

A.2.3. Critically Spaced Array

For a critically spaced array, with $d_y = d_z = \lambda/2$, the phase gradients from (A.27) and (A.28) simplify to

$$\Delta\gamma_y = -\pi \sin \phi \cos \theta + n_y 2\pi \quad (\text{A.29})$$

$$\Delta\gamma_z = -\pi \sin \theta + n_z 2\pi. \quad (\text{A.30})$$

Consequently, in order to identify the steering angles based on the applied linear phase gradients, the elevation angle could be calculated first based on (A.30) as

$$\theta = \arcsin\left(-\frac{\Delta\gamma_z - n_z 2\pi}{\pi}\right). \quad (\text{A.31})$$

And using the result of (A.31), consequently the azimuth angle could be calculated based on (A.29) as

$$\phi = \arcsin\left(-\frac{\Delta\gamma_y - n_y 2\pi}{\pi \cos \theta}\right). \quad (\text{A.32})$$

A.3. Beamforming

In case of a fully digital beamforming, the precoding weights forming a precoding vector \mathbf{w} are calculated based on the estimate of the channel response \mathbf{h} . In Massive MIMO literature simple linear precoding schemes, like maximum ratio (MR), zero forcing (ZF), and regularized zero forcing (RZF) precoding, are considered.

In contrast to the beam steering, using linear phase gradients and uniform excitation amplitudes, covered in Section A.1 and A.2, beamforming provides more control over the radiation pattern. Generally, the magnitude of the precoding vector \mathbf{w}_k can be used to modify the shape of the radiation pattern, while its phase $\angle \mathbf{w}_k$ mainly “controls the scanning capabilities of the array” [74, Sec. 16.7.2].

Note that both the channel and precoding vectors can in a general case be different over different SCs, leading to different radiation pattern for each of the SCs. However, they are equal over SCs for the case of frequency-flat fading channels.

The signal is received by the UE via the effective scalar channel $g_k = \mathbf{h}_k^T \mathbf{w}_k$, where $\mathbf{h}_k^T \in \mathbb{C}^{1 \times M}$ is the channel response vector for the k -th UE and $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the precoding vector for the k -th UE.

The channel vector depends on the relation between the positions of the antennas, on both the base station and the UE side, as well as any potential objects leading to a multi-path propagation environment. For a free-space LoS transmission from a uniform array there is a

certain structure of the channel that is related with the AF definitions in Sections A.1.2 and A.2.2, for horizontally-steered ULA and for URA, respectively.

For the k -th UE located at an azimuth observation angle ϕ_k receiving signal from a linear array, the elements of the channel response vector are essentially corresponding to complex conjugates of the elements of the sum that constitute $\text{AF}(\psi(\phi_k))$ from (A.7) when no steering is applied (i.e., $\Delta\gamma = 0$, in (A.8)). That is, adapted from [1, Eq. 1.23]:

$$\mathbf{h}_k^T = \sqrt{\beta_k} [1, e^{j\kappa d \sin \phi_k}, \dots, e^{j(M-1)\kappa d \sin \phi_k}] \quad (\text{A.33})$$

where κ is the angular wavenumber, d is the spacing between antennas, M is the number of antennas, and β_k is the average channel gain for the k -th UE.

For MR precoding, the precoding vector $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is calculated per UE, as a complex conjugate of the channel response vector (assuming that a perfect channel knowledge is available at the base station side – otherwise an estimate of the channel response needs to be used instead), as follows (see [1, Eq. 1.48]):

$$\mathbf{w}_k = \rho_k \mathbf{h}_k^* \quad (\text{A.34})$$

where ρ_k is the k -th UE power normalization factor.

In case of a free-space LoS channel, the phase of the precoding vector $\angle \mathbf{w}_k$ maps to a linear phase progression using a phase gradient $\Delta\gamma_k$ selected to maximize the radiation towards ϕ_k :

$$\angle \mathbf{w}_k = \Delta\gamma_k [0, \dots, (M-1)]^T. \quad (\text{A.35})$$

For ZF, the precoding matrix \mathbf{W} is calculated as the pseudoinverse of the channel matrix \mathbf{H} , to address the inter-user interference (IUI):

$$\mathbf{W} = \rho \mathbf{H}^+ = \rho \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1} \quad (\text{A.36})$$

where $\mathbf{H}^H \in \mathbb{C}^{M \times K}$ is the conjugate transpose of the channel matrix $\mathbf{H} \in \mathbb{C}^{K \times M}$ and ρ provides power normalization.

Appendix B

Intermodulation Beams

B.1. Intermodulation Products

Each of the TX chains, providing input signals to the antennas, is assumed to contain nonlinear processing elements. Both in the digital part of the processing chain (PAPR reduction using peak-limiter) and in the analog part of the chain (PA, which is assumed to be linearized up to the saturation level).

A memoryless polynomial model, representing the nonlinear processing, is considered:

$$x' = \mu_0 + \mu_1 x^1 + \dots + \mu_n x^n + \dots + \mu_N x^N \quad (\text{B.1})$$

where ($n = 0, \dots, N$), N is the order of the model, x is the input signal, x' is the output signal, μ_n is the n -th order coefficient.

Focusing on the odd-order components of a third-order system, (B.1) simplifies to

$$x' = \mu_1 x^1 + \mu_3 x^3. \quad (\text{B.2})$$

In order to characterize a third-order system from (B.2), a three-tone test is required. A following sinusoidal input is assumed:

$$x(t) = \alpha_1 \sin(\omega_1 t + \gamma_1) + \alpha_2 \sin(\omega_2 t + \gamma_2) + \alpha_3 \sin(\omega_3 t + \gamma_3) \quad (\text{B.3})$$

where ω_i is the angular frequency, α_i is the amplitude, and γ_i is the phase of the i -th tone.

It is clear that $\mu_1 x^1$ component in (B.2) will provide a linearly scaled version of the input signal x , without any new frequency components and no phase shifts. Hence this part can be ignored when studying the third-order intermodulation products (IMPs).

Focusing on the $\mu_3 x^3$ part of the model in (B.2), plugging in the input signal of (B.3), and assuming $\mu_3 = 1$ for brevity, one gets:

$$\begin{aligned}
x(t)^3 &= (\alpha_1 \sin(\omega_1 t + \gamma_1) + \alpha_2 \sin(\omega_2 t + \gamma_2) + \alpha_3 \sin(\omega_3 t + \gamma_3))^3 \\
&= \alpha_1^3 \sin^3(\omega_1 t + \gamma_1) + \alpha_2^3 \sin^3(\omega_2 t + \gamma_2) + \alpha_3^3 \sin^3(\omega_3 t + \gamma_3) \\
&\quad + 3\alpha_1^2 \sin^2(\omega_1 t + \gamma_1) (\alpha_2 \sin(\omega_2 t + \gamma_2) + \alpha_3 \sin(\omega_3 t + \gamma_3)) \\
&\quad + 3\alpha_2^2 \sin^2(\omega_2 t + \gamma_2) (\alpha_1 \sin(\omega_1 t + \gamma_1) + \alpha_3 \sin(\omega_3 t + \gamma_3)) \\
&\quad + 3\alpha_3^2 \sin^2(\omega_3 t + \gamma_3) (\alpha_1 \sin(\omega_1 t + \gamma_1) + \alpha_2 \sin(\omega_2 t + \gamma_2)) \\
&\quad + 6\alpha_1 \alpha_2 \alpha_3 \sin(\omega_1 t + \gamma_1) \sin(\omega_2 t + \gamma_2) \sin(\omega_3 t + \gamma_3). \tag{B.4}
\end{aligned}$$

The first group of factors in (B.4) of the form $\sin^3(\varphi)$ provides components at the fundamental and the third harmonic frequencies, as can be shown using the following trigonometric identity:

$$\sin^3(\varphi) = \frac{3 \sin(\varphi) - \sin(3\varphi)}{4}. \tag{B.5}$$

The second group of the factors in (B.4) of the form $\sin^2(\varphi_A) (\sin(\varphi_B) + \sin(\varphi_C))$, where $A, B, C = (1, 2, 3)$ and $A \neq B \neq C$, can be reformulated using trigonometric identities in the following way:

$$\begin{aligned}
&3\alpha_A^2 \sin^2(\varphi_A) (\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C)) = \\
&= \frac{3}{2} \alpha_A^2 ((\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C)) - \cos(2\varphi_A) (\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C))) \\
&= \frac{3}{2} \alpha_A^2 (\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C)) \\
&\quad - \frac{3}{2} \alpha_A^2 (\alpha_B \cos(2\varphi_A) \sin(\varphi_B) + \alpha_C \cos(2\varphi_A) \sin(\varphi_C)) \\
&= \frac{3}{2} \alpha_A^2 (\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C)) \\
&\quad - \frac{3}{4} \alpha_A^2 \alpha_B (\sin(2\varphi_A + \varphi_B) - \sin(2\varphi_A - \varphi_B)) \\
&\quad - \frac{3}{4} \alpha_A^2 \alpha_C (\sin(2\varphi_A + \varphi_C) - \sin(2\varphi_A - \varphi_C)) \\
&= \frac{3}{2} \alpha_A^2 (\alpha_B \sin(\varphi_B) + \alpha_C \sin(\varphi_C)) \\
&\quad + \frac{3}{4} \alpha_A^2 \alpha_B (\sin(2\varphi_A - \varphi_B) - \sin(2\varphi_A + \varphi_B)) \\
&\quad + \frac{3}{4} \alpha_A^2 \alpha_C (\sin(2\varphi_A - \varphi_C) - \sin(2\varphi_A + \varphi_C)) \tag{B.6}
\end{aligned}$$

It follows that this group of factors, expanded in (B.6), has components at the fundamental frequencies and IMPs of $2A+B$ and $2A-B$ types.

Finally, the third group of factors in (B.4) of the form $\sin(\varphi_A) \sin(\varphi_B) \sin(\varphi_C)$, can be reformulated using trigonometric identities in the following way:

$$\begin{aligned}
& 6\alpha_A\alpha_B\alpha_C \sin(\varphi_A) \sin(\varphi_B) \sin(\varphi_C) = \\
& = \frac{6}{2}\alpha_A\alpha_B\alpha_C \sin(\varphi_A) (\cos(\varphi_B - \varphi_C) - \cos(\varphi_B + \varphi_C)) \\
& = \frac{6}{2}\alpha_A\alpha_B\alpha_C (\sin(\varphi_A) \cos(\varphi_B - \varphi_C) - \sin(\varphi_A) \cos(\varphi_B + \varphi_C)) \\
& = \frac{3}{2}\alpha_A\alpha_B\alpha_C (\sin(\varphi_A + \varphi_B - \varphi_C) + \sin(\varphi_A - \varphi_B + \varphi_C) - \sin(\varphi_A - \varphi_B - \varphi_C) \\
& \quad - \sin(\varphi_A + \varphi_B + \varphi_C)) \\
& = \frac{3}{2}\alpha_A\alpha_B\alpha_C (\sin(\varphi_A + \varphi_B - \varphi_C) + \sin(\varphi_A - \varphi_B + \varphi_C) + \sin(-\varphi_A + \varphi_B + \varphi_C) \\
& \quad - \sin(\varphi_A + \varphi_B + \varphi_C)) \tag{B.7}
\end{aligned}$$

Apart from components of the A+B–C type, there is also A+B+C factor in (B.7) which can be ignored when considering the in-band impact of IMPs in an OFDM signal, since it leads to a frequency that is far away from the in-band (i.e., closer to the range of 3rd harmonics).

In summary, all the third-order IMPs of the 2A–B and A+B–C types resulting from (B.4) are as given in Table B.1. There are up to 6 frequencies related with 2A–B type of IMPs and up to 3 frequencies related with A+B–C type of IMPs.

Table B.1: Intermodulation products (IMP) from (B.4), grouped in 2A–B and A+B–C classes.

2A–B	A+B–C
$\frac{3}{4}\alpha_1^2\alpha_2 \sin(2\omega_1t - \omega_2t + 2\gamma_1 - \gamma_2)$	$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin(\omega_1t + \omega_2t - \omega_3t + \gamma_1 + \gamma_2 - \gamma_3)$
$\frac{3}{4}\alpha_1^2\alpha_3 \sin(2\omega_1t - \omega_3t + 2\gamma_1 - \gamma_3)$	$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin(\omega_1t + \omega_3t - \omega_2t + \gamma_1 + \gamma_3 - \gamma_2)$
$\frac{3}{4}\alpha_2^2\alpha_1 \sin(2\omega_2t - \omega_1t + 2\gamma_2 - \gamma_1)$	$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin(\omega_2t + \omega_3t - \omega_1t + \gamma_2 + \gamma_3 - \gamma_1)$
$\frac{3}{4}\alpha_2^2\alpha_3 \sin(2\omega_2t - \omega_3t + 2\gamma_2 - \gamma_3)$	
$\frac{3}{4}\alpha_3^2\alpha_1 \sin(2\omega_3t - \omega_1t + 2\gamma_3 - \gamma_1)$	
$\frac{3}{4}\alpha_3^2\alpha_2 \sin(2\omega_3t - \omega_2t + 2\gamma_3 - \gamma_2)$	

Assuming an equal amplitude of each of the input tones $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$, the sum power in the two classes of IMPs is the same. However, for the A+B–C class it is concentrated in 50% less IMP tones. Hence it is expected that the power of A+B–C tones will be 6 dB higher than 2A–B.

Such a nonlinear system described by (B.2), introduces new spectral components, that were not part of the original input signal x . For a single-tone input x at f_1 frequency, the third-order harmonic distortion is introduced at $3f_1$ frequency, as visible in the example provided

in Figure B.1(a). With a multi-tone input, apart from the harmonic distortion, there is also the intermodulation distortion (IMD) introduced. That IMD consists of intermodulation products (IMPs) located at frequencies related with the fundamental frequencies of the input signal, which is visualized for 2 tones in Figure B.1(b) and 3 input tones in Figure B.1(c).

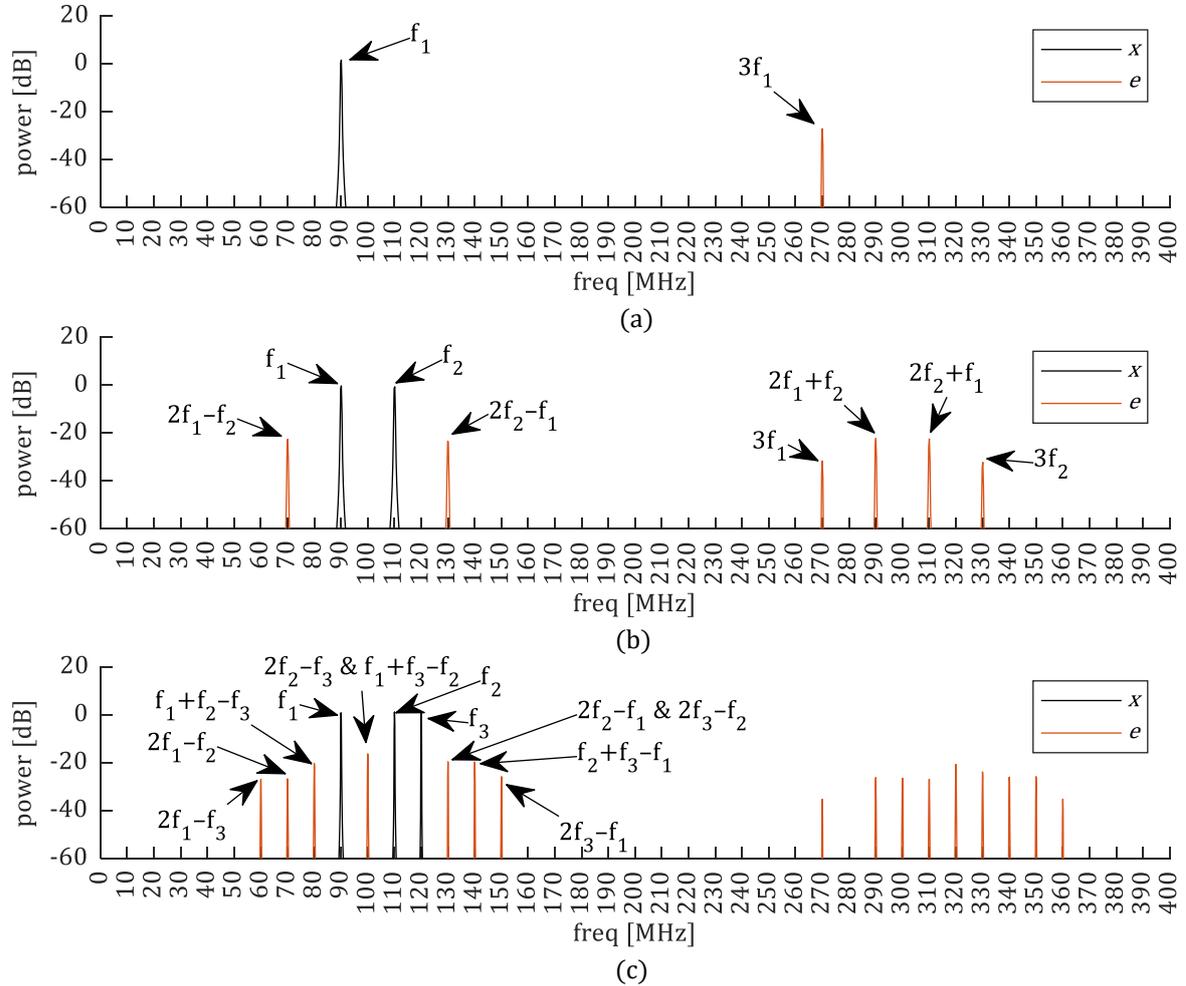


Figure B.1: Power spectrum for the passband representation of the transmitted signal, with 1 to 3 input tones passed through a third-order nonlinearity. Input signal x and error e (difference between equalized output and the input) is visualized.

It can be noticed in Figure B.1(c) as well as in Table B.2, which lists the exact frequencies of the IMP tones, that the IMPs fall at frequencies that are different than the fundamental frequencies in the input signal x . It is clear from Table B.2 that 100 MHz is expected to get most of the IMP power, which can be visibly confirmed in Figure B.1(c).

Table B.2: List of the frequencies of the IMP tones for the 3-tone input from Figure B.1(c).

2A–B	A+B–C
$2f_1 - f_3 = 60$ MHz	$f_1 + f_2 - f_3 = f_2 + f_1 - f_3 = 80$ MHz $f_1 + f_3 - f_2 = f_3 + f_1 - f_2 = 100$ MHz $f_2 + f_3 - f_1 = f_3 + f_2 - f_1 = 140$ MHz
$2f_1 - f_2 = 70$ MHz	
$2f_2 - f_3 = 100$ MHz	
$2f_2 - f_1 = 130$ MHz	
$2f_3 - f_2 = 130$ MHz	
$2f_3 - f_1 = 150$ MHz	

While in the example of Table B.2, the selection of the fundamental frequencies led to IMP tones at different distinct frequencies, this is not necessarily the case generally. Either with more than 3 tones or with equidistant spacing of 3 tones, some of the IMPs will fall at the fundamental frequencies. When the fundamental tones are equidistantly spaced (i.e., $\omega_1 = \omega_2 - \Delta\omega$ and $\omega_3 = \omega_2 + \Delta\omega$), it follows that some of the frequencies will repeat, as demonstrated in Table B.3.

Table B.3: List of IMPs for a 3-tone input with equidistant frequency spacing. Rows sorted with increasing frequency.

frequency	2A–B factor	A+B–C factor
$\omega_1 - 2\Delta\omega$	$\frac{3}{4}\alpha_1^2\alpha_3 \sin((\omega_2 - 3\Delta\omega)t + 2\gamma_1 - \gamma_3)$	
$\omega_1 - \Delta\omega$	$\frac{3}{4}\alpha_1^2\alpha_2 \sin((\omega_2 - 2\Delta\omega)t + 2\gamma_1 - \gamma_2)$	$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin((\omega_2 - 2\Delta\omega)t + \gamma_1 + \gamma_2 - \gamma_3)$
ω_1	$\frac{3}{4}\alpha_2^2\alpha_3 \sin((\omega_2 - \Delta\omega)t + 2\gamma_2 - \gamma_3)$	
ω_2		$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin(\omega_2 t + \gamma_1 + \gamma_3 - \gamma_2)$
ω_3	$\frac{3}{4}\alpha_2^2\alpha_1 \sin((\omega_2 + \Delta\omega)t + 2\gamma_2 - \gamma_1)$	
$\omega_3 + \Delta\omega$	$\frac{3}{4}\alpha_3^2\alpha_2 \sin((\omega_2 + 2\Delta\omega)t + 2\gamma_3 - \gamma_2)$	$\frac{3}{2}\alpha_1\alpha_2\alpha_3 \sin((\omega_2 + 2\Delta\omega)t + \gamma_2 + \gamma_3 - \gamma_1)$
$\omega_3 + 2\Delta\omega$	$\frac{3}{4}\alpha_3^2\alpha_1 \sin((\omega_2 + 3\Delta\omega)t + 2\gamma_3 - \gamma_1)$	

Instead of selecting the 120 MHz as the 3rd fundamental tone in Figure B.1(c), a different tone, leading to equidistant spacing in frequency could be selected, (e.g., 130 MHz which is visible as $2f_2 - f_1$ in Table B.2). A comparison of these two sets of 3 fundamental tones is visualized in Figure B.2. Clearly, no IMPs fall at the fundamental frequencies in Figure B.2(a), while there are some IMPs falling at each of the fundamental frequency components in Figure B.2(b).

This is very relevant from the perspective of the structure of an OFDM signal, which contains J equally spaced SCs. In this case it is expected that multiple third-order IMPs fall at each of the SCs (i.e., in-band) as well as immediately outside of the used channel (i.e., out-of-band), as demonstrated in Figure B.2(c).

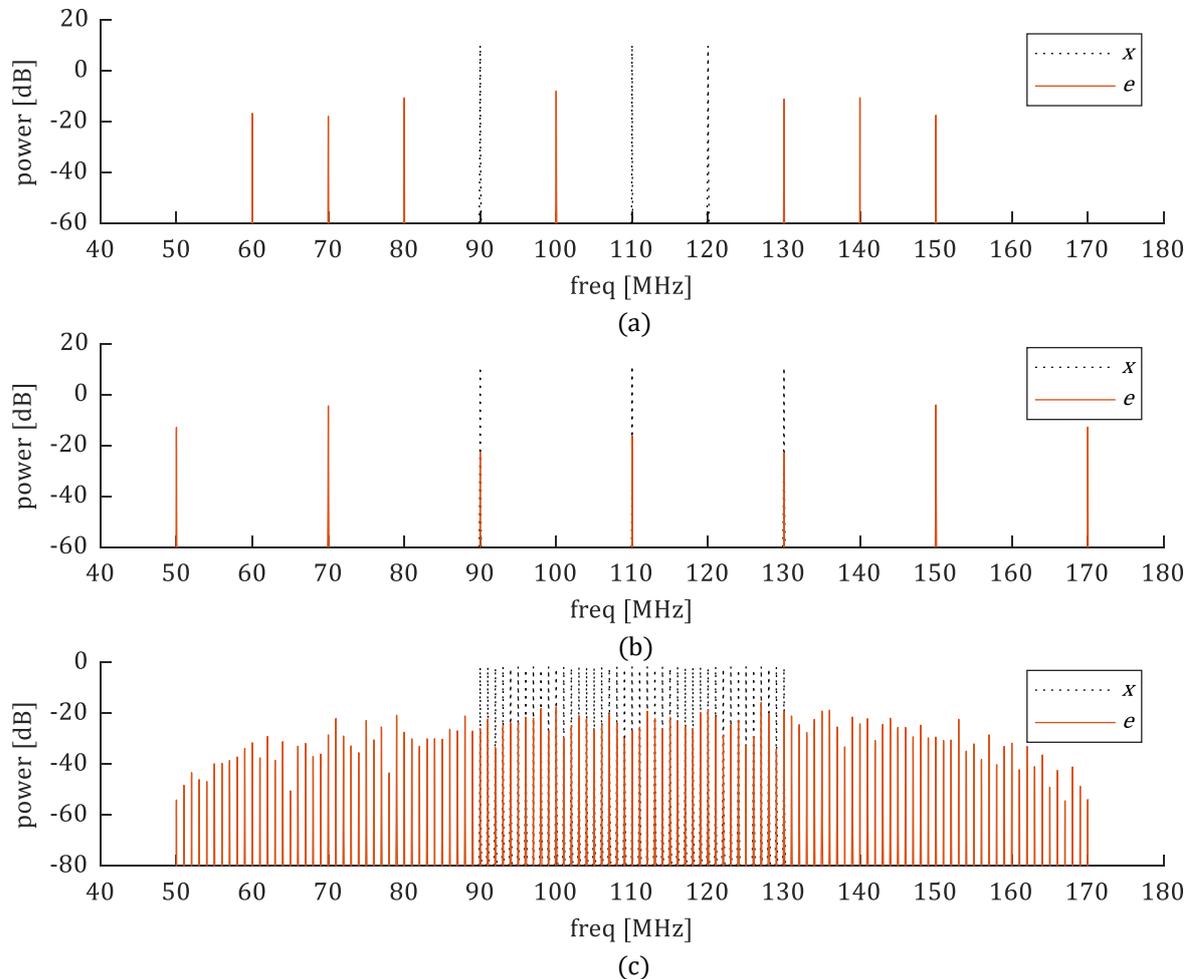


Figure B.2: Power spectrum for the passband representation of the transmitted signal, for the input signal x and error signal e (difference between equalized output and the input) for: (a) x with 90, 110 and 120 MHz tones; (b) x with 90, 110 and 130 MHz tones; (c) a multi-tone signal, covering 40 MHz, with 1 MHz equidistant spacing between tones.

B.2. Intermodulation Distortion Radiated from Antenna Array

Combining the information about the AF from sections A.1.2 and A.2.2 with the information on IMPs from section B.1, one could judge the impact of IMPs on the radiation pattern of IMPs. While third-order IMPs are covered here, this kind of analysis is valid for higher-order IMPs and could be done in analogous way.

The phase differences between antennas in the array determine the beam-pointing directions. For coherent combining with uniform arrays there are certain phase gradients,

over horizontal and vertical antennas, that are applied for beam steering to certain directions in space, as explained in Appendix A.

In the more general case of beamforming (or precoding), the shape of the beam can be altered by using phases that are deviating from the linear phase gradient. The radiation directions for each of the K layers are determined by the precoding vectors \mathbf{w}_k , as explained in Section A.3. For a transmission in LoS channel using MR precoding, phase gradients are used, while with ZF this can deviate.

In MU-MIMO transmission to K single antenna UEs, without FDM and with full allocation of SCs to UEs, there are K MIMO layers for each of the J subcarriers of the transmitted OFDM signal. In principle, when multi-user OFDM signal (with a single beam direction per layer applicable to all SCs) is passed through a nonlinearity there are two classes of generated IMPs that can be distinguished [22]:

1. explicitly user-directed, which are beamformed in the same directions as the users:
 - a. for 2A–B: when both A and B are components from the same MIMO layer;
 - b. for A+B–C: when either A or B or both are from the same MIMO layer as C;
2. spatially filtered, which are implicitly beamformed in distinct directions (resulting from the precoding weights applied to the data signals):
 - a. for 2A–B: when A and B are components from different MIMO layers;
 - b. for A+B–C: when C is from different MIMO layer than A and B.

Clearly, the IMPs in the explicitly user-directed category radiate to the directions of the MIMO layers existing in the input signal, since:

- when $\Delta\gamma_A = \Delta\gamma_B = \Delta\gamma_k$ then $(2\Delta\gamma_A - \Delta\gamma_B) = \Delta\gamma_k$, which radiates in the same direction as the k -th MIMO layer;
- when $\Delta\gamma_A = \Delta\gamma_C = \Delta\gamma_k$ then $(\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C) = \Delta\gamma_B$ and when $\Delta\gamma_B = \Delta\gamma_C = \Delta\gamma_k$ then $(\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C) = \Delta\gamma_A$, which results in radiation to some of the directions existing in the original signal.

While the spatially filtered IMPs radiate to directions resulting from linear combinations of the per antenna phases in the original signal.

B.2.1. Example for Uniform Linear Array

For example, assuming ULA with half-wavelength spacing, 3 UEs placed at azimuth angles of -40° , 5° , and 30° the spatially filtered IMPs are expected to radiate in 9 distinct directions, different than the directions of the UEs, as given in Table B.4 and visualized in Figure B.3.

Table B.4: Summary of phase gradients and resulting steering angles of the spatially filtered intermodulation products (IMPs) for UEs placed at azimuth angles of -40° , 5° , and 30° .

IMP type	spatially filtered IMP	phase gradient [deg]	steering angle [deg]
2A-B	$2\Delta\gamma_1 - \Delta\gamma_2$	247.1	38.8
	$2\Delta\gamma_1 - \Delta\gamma_3$	321.4	12.4
	$2\Delta\gamma_2 - \Delta\gamma_1$	-147.1	54.8
	$2\Delta\gamma_2 - \Delta\gamma_3$	58.6	-19.0
	$2\Delta\gamma_3 - \Delta\gamma_1$	-295.7	-20.9
	$2\Delta\gamma_3 - \Delta\gamma_2$	-164.3	65.9
A+B-C	$\Delta\gamma_1 + \Delta\gamma_2 - \Delta\gamma_3$	190.0	70.8
	$\Delta\gamma_1 + \Delta\gamma_3 - \Delta\gamma_2$	41.4	-13.3
	$\Delta\gamma_2 + \Delta\gamma_3 - \Delta\gamma_1$	-221.4	-50.4

Note that A+B-C type of IMPs with A and B from the same MIMO layer are not separately shown in Table B.4, since they map to the same phase gradients as corresponding 2A-B.

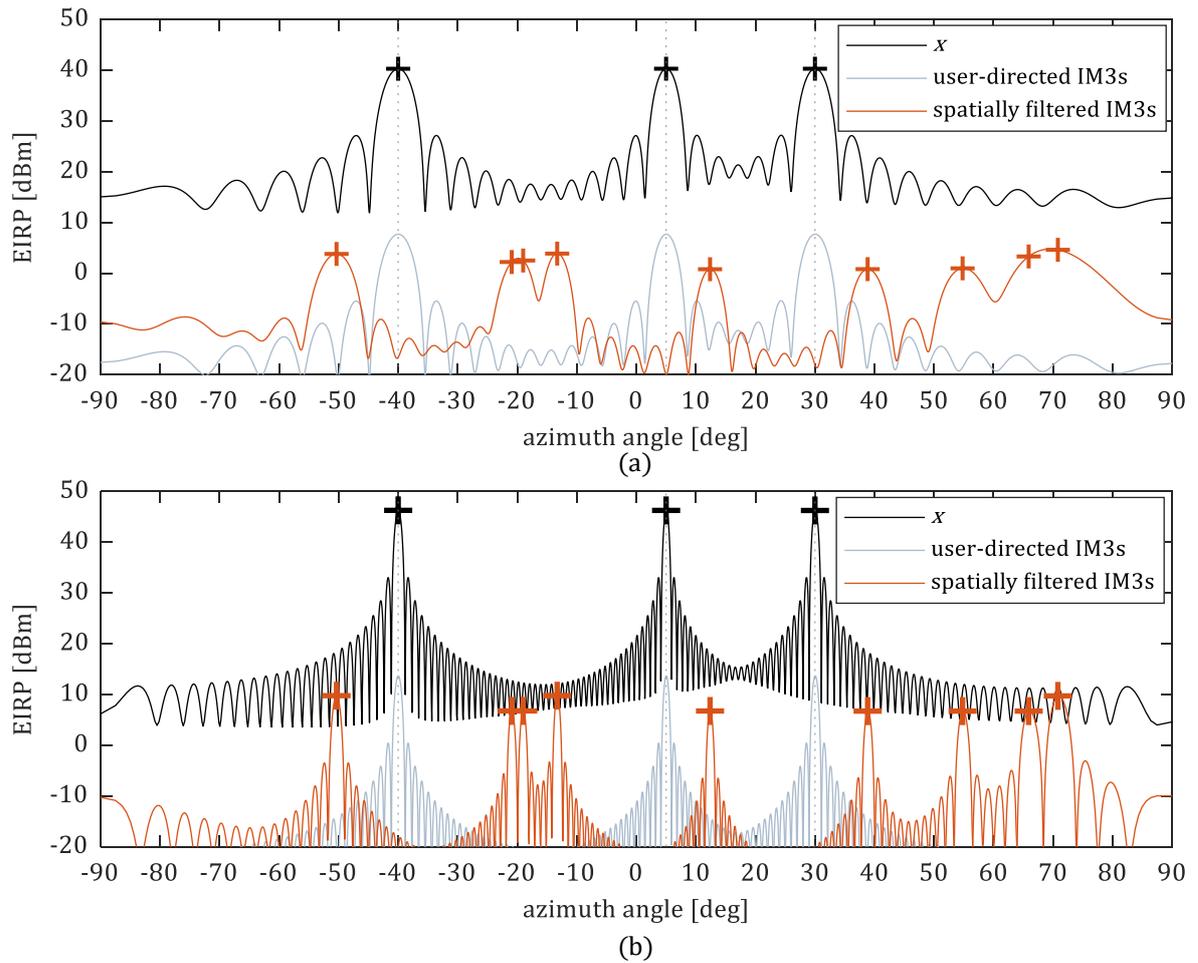


Figure B.3: Radiation patterns of wanted signal x , as well as user-directed and spatially filtered third-order intermodulation products (IM3s): (a) 32 antennas; (b) 128 antennas.

By comparing Figure B.3(a) and Figure B.3(b), it is apparent that that when the IM beams are angularly relatively close to each other it may not be possible to judge that from the radiation pattern plot alone.

While the steering angles selected in the above example led to IMPs with distinct steering angles, not existing in the set of fundamental signals, it is not the case generally. It is not excluded that spatially filtered IMP leads in practice to the same phase as in the set of fundamental phase gradients Γ as well. This happens when either $(2\Delta\gamma_A - \Delta\gamma_B + n2\pi) \in \Gamma$ or $(\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C + n2\pi) \in \Gamma$, for any integer n .

For example, considering 3 UEs a situation that IMP generated by interaction of the 1st and 2nd UE lead to IMP falling at the 3rd UE location happens when either

$$(2\Delta\gamma_A - \Delta\gamma_B + n2\pi) = \Delta\gamma_C \quad (\text{B.8})$$

$$(\Delta\gamma_A + \Delta\gamma_B - \Delta\gamma_C + n2\pi) = \Delta\gamma_C. \quad (\text{B.9})$$

This leads to

$$\Delta\gamma_C = (2\Delta\gamma_A - \Delta\gamma_B + n2\pi) \quad (\text{B.10})$$

$$\Delta\gamma_C = \frac{(\Delta\gamma_A + \Delta\gamma_B + n2\pi)}{2}. \quad (\text{B.11})$$

It is straightforward to identify the angles at which the 3rd UE should be positioned to receive the 2A–B type of IMPs from the other beams, since (B.10) does not depend on the 3rd UE position. For example, assuming $\phi_1 = -20^\circ$ and $\phi_2 = 0^\circ$ the related phase gradients are $\Delta\gamma_1 \approx 61.6^\circ$ and $\Delta\gamma_2 = 0^\circ$, which leads to IMPs with $\Delta\gamma_{d1} \approx 123.1^\circ$ and $\Delta\gamma_{d2} \approx -61.6^\circ$, leading to $\phi_{d1} \approx -43.1^\circ$ and $\phi_{d2} = 20^\circ$. The related radiation patterns for this scenario are visualized in Figure B.4. Clearly, positioning the 3rd UE at either of the angles with IMPs will lead to radiation of that spatially filtered IMP towards that UE.

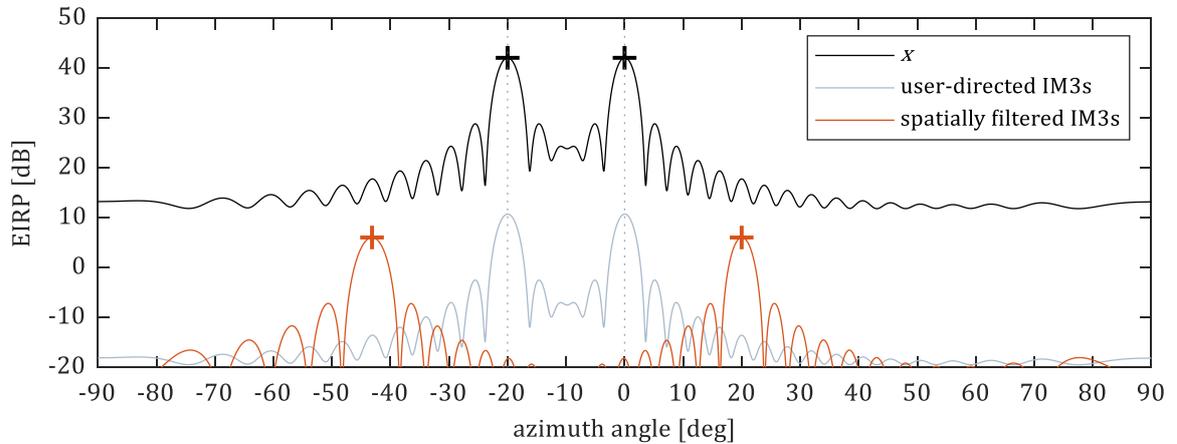


Figure B.4: Radiation patterns of wanted signal x , as well as user-directed and spatially filtered third-order IMPs (IM3s).

What may not be apparent at first glance, at least not by judging the 2A–B type of IMPs only, is the other angular positions of the 3rd beam that lead to a similar situation of having IMPs transmitting to that 3rd beam direction. Different combinations of such locations, based on (B.11) are provided in Table B.5.

Table B.5: Phase gradients and steering angles for identified intermodulation beams.

phase gradient type	n	phase gradient [deg]	steering angle [deg]
$\Delta\gamma_C = \frac{(\Delta\gamma_A + \Delta\gamma_B + n2\pi)}{2}$	0	30.8	-9.8
	1	210.8	56.0
	-1	-149.2	56.0

The first angle of about -10 deg could likely be expected, leading to equidistant spacing between the phase gradients. All the identified locations for the 3rd beam based on both phase gradient types are presented in Figure B.5.

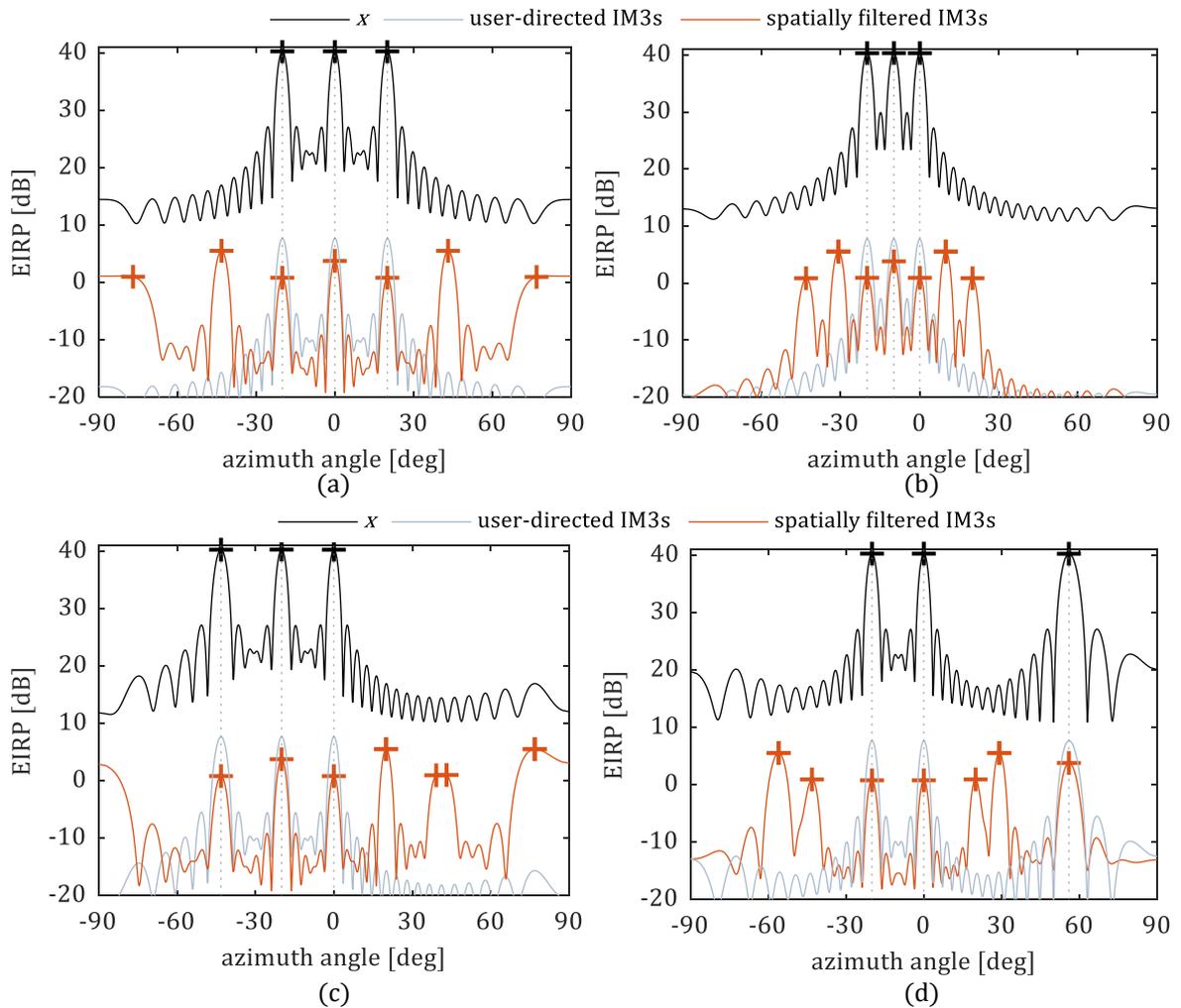


Figure B.5: Radiation patterns, for linear array with 32 antennas, for all the identified unfavorable locations of adding the 3rd beam.

Appendix C

The Link-Level Simulation Environment

A proprietary link-level simulation (LLS) environment was developed by the PhD candidate to support realization of the research objectives of this dissertation. The building blocks available in academia- and industry-proven Matlab toolboxes (primarily in 5G Toolbox, Phased Array System Toolbox, and Communications Toolbox) were used to reflect the relevant elements of the downlink (DL) transmission from Massive MIMO base station to multiple UEs over spatially correlated channels.

An overview of the structure and principles of operation of this LLS is provided in Section C.1. While calibration of the performance results is described in Section C.2.

C.1. Structure and Operation

Starting in the frequency domain, the randomly generated bitstreams (reflecting user payload data) are passed through Downlink Shared Channel (DL-SCH) encoder (*nrDLSCH*) and Physical Downlink Shared Channel (PDSCH) modulator (*nrPDSCH*). The modulation symbols are mapped to REs together with PDSCH Demodulation Reference Signal (DMRS) symbols (*nrPDSCHDMRS*), according to the 5G NR resource grid. Here, only PDSCH data and pilot symbols are used in the transmission. Modulation symbols are passed through a precoding block, which provides per antenna signals that are mapped to the time domain in an OFDM modulator (*nrOFDMModulate*).

In the time domain, the TX signal is optionally interpolated to 122.88 Msps (e.g., from 30.72 Msps at the output of the OFDM modulator for 20 MHz 5G NR carrier). Next it is optionally distorted, either by peak power limitation (in line with the models described in Section 4.2.1) or by adding uncorrelated Gaussian noise (reflecting the effects of quantization, in line with the model described in Section 4.2.2). Before passing the signal through the channel, there is an option to calculate and apply compensation signals, based on the novel solution proposed in Chapter 6.

The multi-antenna signal is passed through a spatially correlated clustered delay line (CDL) channel model (*nrCDLChannel*). The CDL channel is used both with predefined delay profiles, representing 3GPP defined multi-path LoS and NLoS channels, as well as with custom profile (e.g., to reflect a free-space LoS channel, or provide alternative

frequency-selective channels in multi-user operation). The channel objects are configured in a way that the channel fading realization changes at the slot boundary. The characteristic of the antenna array is reflected in *nrCDLChannel* using either *phased.ULA* or *phased.URA* at the TX side. The estimate of the channel response, for calculation of precoding weights, is obtained using *nrPerfectChannelEstimate*.

At the RX side of the channel, *awgn* function is optionally used to reflect the effect of RX noise, before transformation to frequency domain in an OFDM demodulator (*nrOFDMDemodulate*).

The quality of the received frequency domain modulation symbols is judged by calculating EVM, while the achievable data rate is determined by BLER measurement. The same equalization procedure (*nrEqualizeMMSE*) is used for both EVM measurement as well as for PDSCH decoding (*nrPDSCHDecode*).

The developed simulation model is deterministic. While pseudorandom bitstreams per UE are used at the input of the DL processing and fading realizations are generated for each slot, the seed of the random number generator is explicitly controlled, to enable repeatability of simulation runs. This allows for isolating influence of selected aspects (e.g., by using the same bitstreams for testing EVM performance with UEs in different angular locations).

Calibration of the BLER performance was done in order to validate the simulation environment, as explained the following section.

C.2. Calibration

The validity of the simulation results obtained with the utilized LLS environment was confirmed on two levels. First, the behavior from the EVM performance perspective was verified against the expectations based on the Massive MIMO literature. Afterwards, the BLER performance was compared against the publicly available results from other LLS-based studies.

The BLER performance was verified in SISO transmission for CDL-D (LoS) and CDL-C (NLoS), by applying Gaussian noise either on the TX or the RX side of the channel. The BLER results for a range of SNDRs for transmission with 256QAM at 948/1024 code rate presented in Figure C.1 were shared already in [62] (co-authored by the PhD candidate). The level of 10% BLER was obtained for the case of the RX noise, with SNR of about 27.1 dB for LoS and 33.8 dB for NLoS channel. Considering that about 25 dB SNR was required in [68] for 10% BLER in AWGN channel with the same modulation order and code rate, the

performance degradation that could be expected in multi-path channels seem to be roughly in line with the results obtained here.

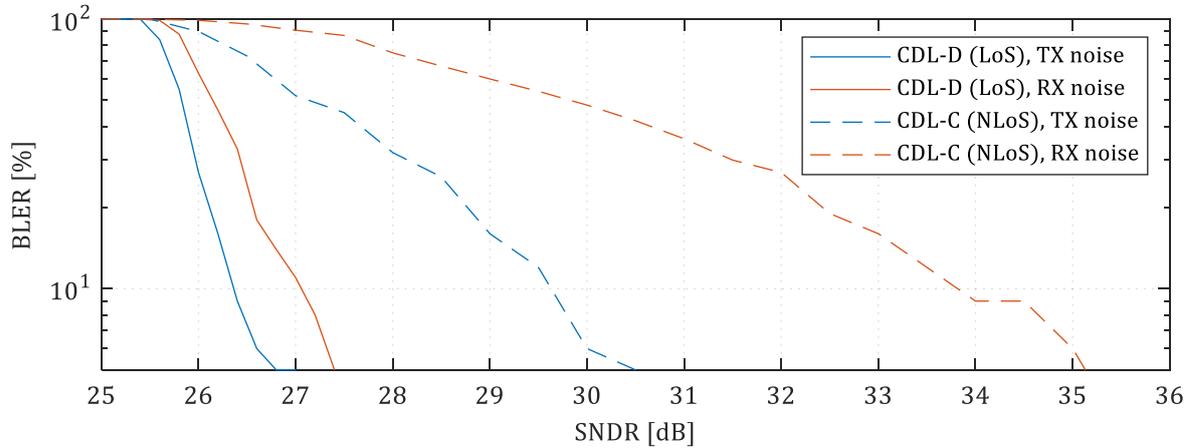


Figure C.1: Comparison of BLER vs wideband signal-to-noise-plus-distortion ratio (SNDR) for SISO transmissions in CDL-C (NLoS) and CDL-D (LoS) channels. Results for Gaussian noise added either on TX or RX side. Based on a figure originally published in [62] (co-authored by the PhD candidate).

To confirm that the BLER performance scales as expected depending on the spectral efficiency, transmission in the CDL-D (LoS) channel with noise added at the TX side of the channel was simulated for 256QAM and 64QAM with relatively high and relatively low code rates. The results, presented in Table C.1, clearly show that the level of required SNDR (hence also tolerable EVM) strongly depends on the used modulation order and code rate. It is visible that roughly the same SNDR is required for 64QAM with the highest possible code rate (i.e., 948/1024) like for 256QAM with the lowest possible code rate (i.e., 682.5/1024).

Table C.1: SNDR for 10% BLER for SISO transmission using 256QAM and 64QAM at different code rates.

	SNDR for 10% BLER [dB]			
	256QAM 948/1024	256QAM 682.5/1024	64QAM 948/1024	64QAM 466/1024
Own LLS, for CDL-D (LoS)	26.4	19.7	19.9	11.2
From [68], for AWGN channel	25.3	18.8	N/A	10.3

The comparison of the results shared in Figure C.1 and Table C.1 with the results from [68] suggests that the BLER performance obtained with own LLS environment in SISO transmission is in line with expectation.

Appendix D

Simulation Results Supplementing the Core Analysis

The simulation results presented in Chapter 4.7 directly support validation of the first thesis of this dissertation. While the collection of simulation results presented in this appendix covers a handful of supplementary verifications. Some of which were performed to obtain a better understanding of the impact of distortion but were not seen as essential elements supporting the thesis. However, some are provided here to support the claims made in the discussion of the simulation results presented in Chapter 4.7.

There are no direct links between the aspects covered in different sections of this chapter.

D.1. Higher-Order Intermodulation Products

The EVM performance for ICF was compared in Section 4.7.2 with the results of the precoding-based prediction model, relying solely on the third-order IMPs. It was observed that the RMS EVM for the best-case was lower than the prediction based on the best-case IM3s by up to about 0.9 dB. The primary reason for that deviation was assumed to be related with the influence of higher-order IMPs (which are expected to be introduced by ICF). This reasoning is explored here in a bit more detail.

As an example, a comparison of the radiation pattern of the distortion signal d plotted based on a 1 slot long simulation for the best-case scenario for 3 UEs (with the wanted signal x), with the expected radiation patterns of spatially filtered third- and fifth-order IMPs (IM3s and IM5s in short, respectively) calculated based on precoding weights, is presented in Figure D.1.

It is clearly visible that in this case the spatially filtered IMPs, for both third- and fifth-order, are expected to be radiated away from the UEs (i.e., there are no visible lobes at the UE angular locations). Hence what is radiated towards UEs are the explicitly user-directed IMPs and potentially higher-order spatially filtered IMPs. This demonstrates the validity of the reasoning that the EVM prediction using third-order IMPs may be pessimistic.

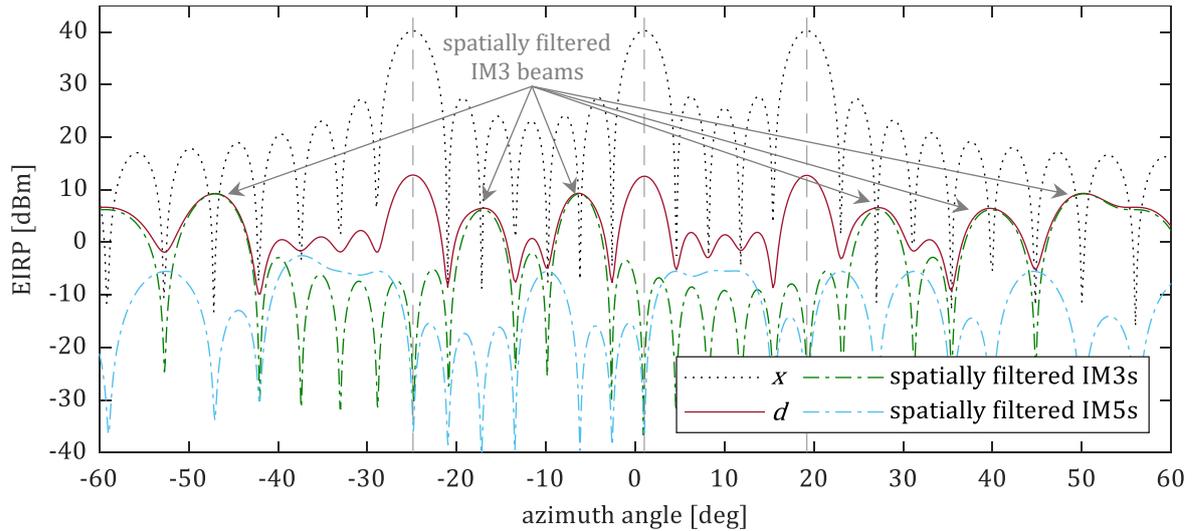


Figure D.1: A comparison of the radiation patterns of the wanted signal x and distortion d , based on link-level simulation results, with the expectations of the third- and fifth-order intermodulation (IM) beams calculated based on precoding weights. UEs are located at the azimuth angles marked by the vertical dashed lines.

D.2. Angular Positions of Users in Non-Line-of-Sight Channels and Power Variation Over Antennas

While spatially multiplexed UEs may create worst-case sets of angular directions leading to lack of improvement in EVM with increasing number of UEs, as demonstrated in Section 4.7.2 for free-space and multi-path LoS scenarios, this kind of behavior is not expected in NLoS channels. This seems to be confirmed by the apparent lack of notable differences in the EVM performance obtained for the best- and worst-cases of LoS angular locations, as visible in the summary provided in Table D.1, when transmission over CDL-A (NLoS) channel was simulated for 100 slots.

Table D.1: RMS over 100 slots of RMS EVMs (over UEs per slot) for CDL-A (NLoS).

Number of UEs (K)	RMS of per slot EVM [dB]				
	1	2	3	4	6
LoS best-case	-24.0	-26.9	-27.4	-27.3	-27.5
LoS worst-case	-23.9	-26.6	-27.1	-27.4	-27.5
Δ (best – worst)	-0.1	-0.3	-0.3	0.1	0.0

Similarly as for the multi-path LoS results, demonstrated for the CDL-D channel in Section 4.7.2, usage of equal power budget per SC led to about 0.2 dB lower EVM, compared to the results presented in Table D.1 for the power budget used to equalize channel gains.

Clearly, a significant difference in the EVM performance between $K = 1$ and $K > 1$ can be observed in the results shared in Table D.1. It is expected that the primary reason for that

about 3 dB gap was the variation of the per antenna norms of precoding weights. Note that it was demonstrated in [2] for 1 UE in single-carrier frequency-selective channel that the variation of power over antennas reduces with increasing delay spread.

In order to investigate whether the power variation over antennas, resulting effectively in different PAPR reduction levels in different antennas, may be the reason for the observed performance gap a verification with peak limiting threshold scaled with the average per antenna power was considered. While this is not exactly equivalent with the approach to the PAPC taken in [48] where the beamforming weights were adjusted to reduce the variation of the mean power over antennas, it shows to some extent what could happen when the same PAPR is targeted in all antennas. However, the difference is that a higher power of distortion is expected to be radiated from antennas with a higher mean power, despite having same PAPR targeted, since the relative power of distortion will be higher. This is expected to influence the radiation pattern of distortion, since it maps to using nonuniform amplitude excitation of the array.

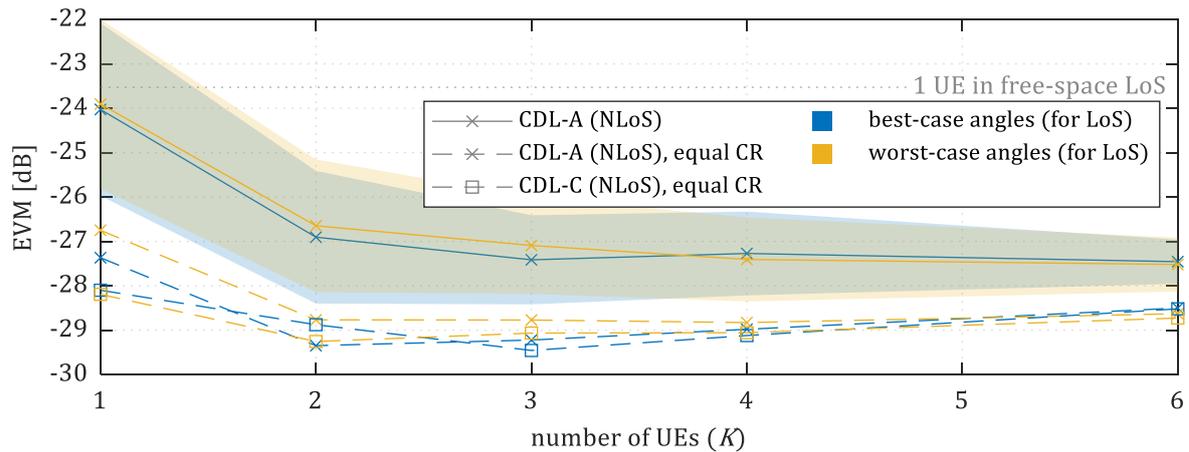


Figure D.2: Comparison of EVM simulation results for NLoS channels: CDL-A (RMS: solid; RMS \pm std dev: filled ranges), as well as CDL-A and CDL-C with clipping ratios (CRs) scaled with per antenna norms of precoding weights, for UEs located at the sets of angles of the best- and worst-case scenarios for LoS.

In order to verify the validity of the assumption that the power variation over antennas significantly influenced the CDL-D (LoS) and CDL-A (NLoS) results, presented separately in Figure 4.19 and Figure D.2, additional verification was performed with CRs adjusted based on the per antenna norms of precoding weights, so that effectively same CR was targeted per antenna. The related improvements of the RMS EVM values, for both CDL-D and CDL-A, are summarized in Table D.2. Clearly, a significant positive impact is visible for NLoS, in both best- and worst-cases. Actually, the resulting EVMs for 2–6 UEs in CDL-A were less than 0.5 dB away from the -29.0 dB level observed for the best-case scenario

for 6 UEs in a free-space LoS channel. These improvements, achieved by adjusting the CRs per antenna (so that the effective CR was 6 dB, like in the free-space LoS case), suggest that limitation of the power variation between antennas could lead to lower EVM. For NLoS, the highest benefits in this case could be expected for a low number of UEs.

Table D.2: Improvement in EVM caused by adjustment of the per antenna CRs.

Number of UEs (K)		Improvement in EVM [dB]				
		1	2	3	4	6
CDL-A (NLoS)	LoS best-case	3.3	2.4	1.8	1.7	1.0
	LoS worst-case	2.8	2.1	1.7	1.4	1.1
CDL-D (LoS)	LoS best-case	0.4	0.4	0.3	0.3	0.5
	LoS worst-case	1.0	0.4	0.5	0.3	1.2

For 1 UE however, the EVM was still visibly higher (more than 1 dB) than for more UEs, despite significant performance improvements of about 3 dB compared to the original results from Figure 4.19. A part of that difference is likely related with the lower number of directions with significant radiation in case of 1 UE, compared to the multi-UE cases. However, it is assumed that this aspect is combined with the angular resolution of radiation, limited by the array aperture. This is considered to be the primary reason for the EVM improving with increasing the number of antennas for a transmission to 1 UE in CDL-C (NLoS) channel, which is especially visible for the range from 1 to 16 antennas in the results published already in [62]. There, the relation of the EVM (resulting from the distortions introduced by HC) and M antennas (verified in the range from 1 to 512) was explored for free-space LoS, CDL-D (LoS), and CDL-C (NLoS) channels. In the published results, visualized in Figure D.3, a considerable improvement (of about 5 dB) for NLoS channel can be observed when the number of antennas M is increased from 1. However, most of that improvement was actually already visible with 16 antennas and no noticeable differences in EVM were observed when increasing the number of antennas beyond 128.

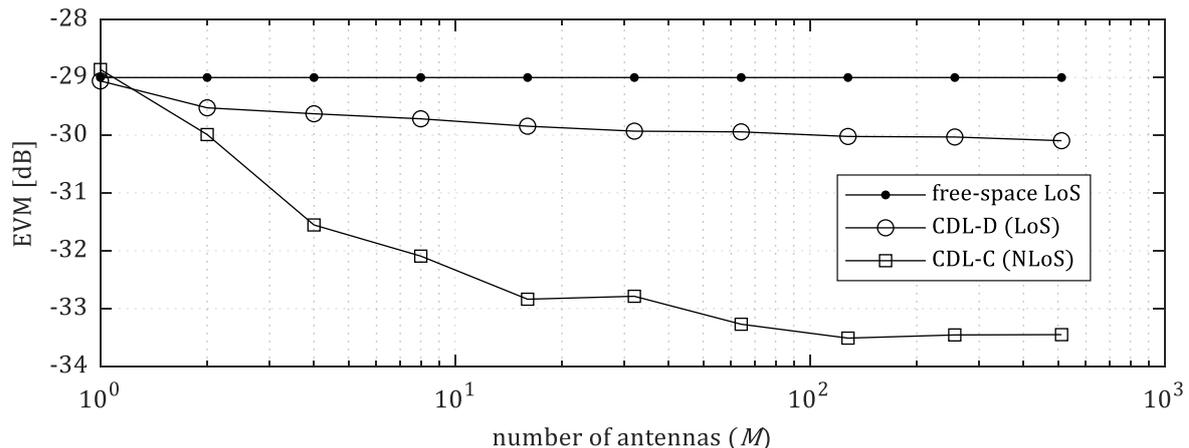


Figure D.3: For 1 UE, when the number of antennas M is increased (from 1 to 512), the EVM resulting from hard clipping (HC) changes differently depending on the channel type. Based on a figure originally published in [62] (co-authored by the PhD candidate).

Apart from the possible EVM performance improvements caused by an increased array aperture (achieved here by increasing the number of antennas in the array), leading to more directional radiation pattern of distortion, the EVM from nonlinear distortion is generally not expected to noticeably improve with increasing number of antennas on the base station side. This is clearly visible in the apparent lack of EVM reduction for 1 UE in a free-space LoS channel. Both for ICF (when the results for 32 antennas from Table 4.5 are compared to the per antenna EVM results from Section 4.7.1) and for HC (as clearly visible in the relation of EVM to the number of antennas for free-space LoS in Figure D.3).

D.3. Statistics of the Error Level for Hard Clipping

An initial verification of the EVM for randomly placed UEs was presented in [62] for HC. It was done for free-space LoS (with UEs placed at randomly selected directions of beams from a uniform DFT-based codebook) and CDL-D (multi-path with LoS) channels. The EVM statistics (i.e., min, median, and max of the CCDFs shared in [62]) of median EVMs calculated over UEs per simulation scenario, summarized in Table D.3, were obtained from simulation of 100 random UE placements. Despite the simulation length of 1 slot does not allow to account for the variation of EVM due to changing input data and impact of different fading realizations, a clear trend of EVM reducing with increasing number of UEs can be observed, especially up to $K = 4$, as expected.

When focusing on the median EVM performance (i.e., calculated over the verified 100 spatial scenarios), comparison of the EVM with $K > 1$ to $K = 1$ for free-space LoS cases shows a reduction in EVM of about 1.6 dB for $K = 2$, 3.8 dB for $K = 4$, and 4.4 dB for $K =$

8. This is consistently aligned with the expectation from the precoding-based prediction model, summarized in Table 4.6, with about 0.4 dB offset (suggesting a somewhat pessimistic expectation).

While a similar comparison for CDL-D reveals that improvement in EVM performance of about 2.4 dB with $K = 2$, 4.2 dB with $K = 4$, and 4.4 dB with $K = 8$ was experienced, the median EVM for 1 UE was in this case about 1.4 dB worse than for the free-space LoS cases. In fact, it was also pointed out in [62] that the per antenna power variation can significantly influence the results for CDL-D simulations, which is likely the reason for a relatively high variation of EVM (over UE locations) experienced for 1 UE in CDL-D, visible in Table D.3. This is aligned with the improvements in the EVM performance that were demonstrated in Appendix Section D.2 for both LoS and NLoS scenarios, when equal CR was targeted. This suggests that when the power variation per antenna is reduced, respecting the PAPC, the EVM performance improvements for multi-path channels with LoS expected when increasing the number of UEs could be smaller than observed in Table D.3.

Table D.3: Overview of the EVM statistics of median EVMs over UEs, based on results in [62].

Number of UEs (K)	free-space LoS			CDL-D (multi-path with LoS)		
	min EVM [dB]	median EVM [dB]	max EVM [dB]	min EVM [dB]	median EVM [dB]	max EVM [dB]
1	-29.0	-29.0	-29.0	-29.4	-27.6	-22.8
2	-30.6	-30.6	-29.2	-31.0	-30.0	-27.7
4	-33.3	-32.8	-31.0	-32.9	-31.8	-28.9
8	-33.7	-33.4	-32.0	-32.6	-32.0	-30.2

While the simulation results shared in Table D.3 suggest that the proposed EVM performance prediction models could be used to provide expectations for the typical EVM performance for linearized PAs, modeled with HC, the efficacy of the proposed models for the limiter-based PAPR reduction, modeled with ICF, is considered in Section 4.7.3.2.

D.4. Quantization Noise Radiated Over Multi-Path Channels

It is assumed that the impact of a subset of residual TX impairments could be relatively well represented by Gaussian noise which is uncorrelated between antennas. The quantization operation could potentially be one of such sources of distortion, as explained in Section 4.5.

While it is relatively straightforward to predict the impact to EVM of the number of antennas M and the number of UEs K for a transmission in free-space LoS channels with equal power allocation between UEs, as explained in Section 4.5, the impact of the number

of UEs on the performance in multi-path channels is explored here based on link-level simulation. The results presented in this section were already shared and discussed in [62], co-authored by the PhD candidate.

The relation between the number of UEs K and the level of EVM, was simulated for ULA with 32 antennas transmitting to UEs in multi-path LoS (based on CDL-D) and NLoS (based on CDL-A) channels. The results for 1 slot long simulations with UEs located at randomly selected azimuth angles are presented in Figure D.4. For the multi-path LoS channel, the observed increase in EVM of $10 \log_{10} K$ (i.e., from -44 dB for 1 UE up to -35 dB for 8 UEs), is aligned with the expectation for a free-space LoS channel. The results obtained for NLoS present a similar relation with the number of UEs, with a somewhat larger degradation of EVM when the number of UEs was increased and visibly larger variation of EVM between UEs.

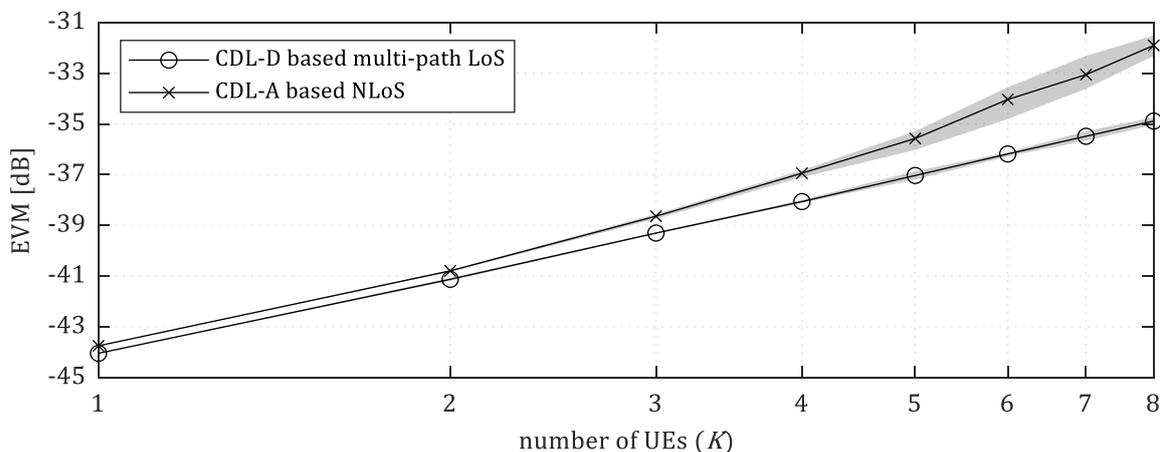


Figure D.4: Examples of the EVM vs the number of UEs K for the quantization noise model. Lines represent the mean EVM over all UEs, while the filled ranges show the EVM variation between UEs. Based on a figure originally published in [62] (co-authored by the PhD candidate).

Clearly, the EVM results obtained for the quantization model with equal power distribution over UEs demonstrate that increasing the number of UEs have a negative effect on the per UE EVM, in both multi-path LoS and NLoS channels. This is in line with the expectations for free-space LoS channels, expressed in Section 4.5. This is a reverse relation compared to nonlinear distortion, as demonstrated in Section 4.7.3.

The impact of an unequal power allocation to UEs (targeting comparable QoS for all UEs) is explored in Section 4.7.5 for both quantization and ICF models.