



Multi-user angle-domain beamforming and non-orthogonal multiple access for millimeter-wave massive MIMO systems

Israa Khaled

► To cite this version:

Israa Khaled. Multi-user angle-domain beamforming and non-orthogonal multiple access for millimeter-wave massive MIMO systems. Networking and Internet Architecture [cs.NI]. Ecole nationale supérieure Mines-Télécom Atlantique; Université Libanaise. Faculté des Sciences (Beyrouth, Liban), 2022. English. NNT: 2022IMTA0294 . tel-03744486

HAL Id: tel-03744486

<https://theses.hal.science/tel-03744486>

Submitted on 3 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPERIEURE MINES-TELECOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE
DÉLIVRÉE CONJOINTEMENT AVEC
L'ÉCOLE DOCTORALE EN SCIENCES ET TECHNOLOGIE -
UNIVERSITE LIBANAISE

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Télécommunications*

Par

Israa KHALED

Multi-user Angle-Domain Beamforming and Non-Orthogonal Multiple Access for Millimeter-Wave Massive MIMO Systems

Formation de voie multi-utilisateurs dans le domaine angulaire et accès
multiple non orthogonal pour les systèmes multi-antennes massifs en ondes
millimétriques

Thèse présentée et soutenue à Brest, le 28/03/2022

Unité de recherche : Lab-STICC

Thèse N° : 2022IMTA0294

Rapporteurs avant soutenance :

Ghaya REKAYA-BEN OTHMAN
Kosai RAOOF

Professeur, Telecom Paris
Professeur, École nationale supérieure d'ingénieurs du Mans (ENSIM)

Composition du Jury :

Président
Examineurs

Kosai RAOOF
Ghaya REKAYA-BEN OTHMAN
Jean-Pierre CANCES
Chadi ABOU-RJEILY
Michel JEZEQUEL
Bachar ELHASSAN
Charlotte LANGLAIS
Ammar EL FALOU

Professeur, ENSIM
Professeur, Telecom Paris
Professeur, Université de Limoges
Professeur, Université Libanaise Américaine
Professeur, IMT Atlantique
Professeur, Université Libanaise
Directrice d'études, IMT Atlantique
Ingénieur de recherche, King Abdullah University of Science
and Technology (KAUST)

To the paradise of my heart, Mum, and my backbone, Dad
To my beloved sisters and brothers
To my beautiful nephews

Acknowledgments

First and foremost, I would like to thank my supervisors Michel JEZEQUEL, Bachar ELHASSAN, Charlotte LANGLAIS, and Ammar EL FALOU for their guidance during my Ph.D. study. Special thanks to Charlotte LANGLAIS and Ammar EL FALOU, who guided me step by step into the research area and advised me with their expertise and insights.

I would like to thank my colleagues in LabSTICC. This includes but is not limited to Rami KLAIMI and Fatima HAMDAR.

I want to thank every person who taught me, even a letter. Life is not easy, and difficulties surround us in all directions. Thanks to everyone who helped us keep going when it's been so hard. I want to thank also my teacher M. Mohammad RATEB AL NABULSI, who taught us how precious our lives are and how powerful we are to do so much good.

My deepest thanks go to my "Sister in the CITY", Israa AL NAZER, for her continued support and encouragement during my Ph.D. trip.

Last but not least, thanks to my beloved family for their tremendous support and endless love.

Abstract

Multi-User Angle-Domain Beamforming and Non-Orthogonal Multiple Access for MmWave Massive MIMO Systems

Israa Khaled

Electronics Department, IMT Atlantique

With the proliferation of the Internet of Things (IoT) and applications requiring high data rates, ultra-reliability, and low latency, next-generation cellular systems are being called upon to set new standards in the mobile market and respond to the exponential increase in wireless data traffic, the explosion in the number of connected devices and the wide variety of use cases and requirements. Massive multiple-input multiple-output (MIMO) system, non-orthogonal multiple access (NOMA) technique, and millimeter-wave (mmWave) bands have emerged as promising candidates for the next-generation cellular systems. The integration of these technologies can significantly enhance the network capacity, manage massive connectivity, and support the quality-of-service (QoS) requirements of the emerging use cases.

In this regard, this thesis aims to address the main challenges of massive MIMO-NOMA systems for mmWave downlink transmissions, namely the complexity of precoding design, hardware constraints, and the complexity and overhead of channel estimation, which result from the deployment of large-scale antenna arrays (LSAAs). In particular, we focus on the angular information, which is considered a promising partial channel state information (CSI) thanks to the high directionality of mmWave channels and since the feedback overhead does not scale with the number of antennas. By exploiting this feature, we adopt a low-complex angle-domain (AD) digital beamforming (DBF) that uses only digital phase shifters and requires only the angular information to steer a beam in the direction of each user. First, we review the different mmWave channel models and simulators, and we adopt, throughout this thesis, the realistic and statistical mmWave channel model and simulator, called NYUSIM, developed by New York University. And, we investigate the spatial behavior and potentiality of the AD-DBF beamformer in mmWave channels and with LSAAs using NYUSIM. As a reference, we consider the conjugate beamforming technique that maximizes the beamforming (BF) gain at the desired user based on a full CSI. We find that AD-DBF is an appealing BF technique trading-off performance, hardware complexity, and channel estimation overhead. Unfortunately, such systems suffer from high inter-user interference when the users are close to each other in the direction. Second, in this context, two- and multi-user AD MIMO-NOMA schemes, leveraging the NOMA scheduling in the power domain, are proposed when dozens of antennas exist at the base station (BS). Both aim to reduce inter-user interference, especially in a congested cell, and improve the system's spectral efficiency. To do so, we first derived a set of angular-based performance metrics, such as inter-user spatial interference, user channel quality, and sum-throughput. These metrics then participate in the design of user clustering and ordering, and power allocation techniques.

Specifically, the user clustering algorithm groups the users with high spatial interference to perform NOMA. A new user ordering strategy is proposed using the angular-based channel quality metric. In addition, we design a power allocation method that maximizes the angular-based sum-throughput. Simulation results show that the proposed multi-user AD MIMO-NOMA efficiently improves the performance of AD-DBF by achieving up to 39% increase in the spectral efficiency when the number of users is close to that of antennas. Simulation results also show that the proposed user ordering strategy outperforms other limited feedback strategies, and the angular-based power allocation allows for efficient successive interference cancellation (SIC). We find that using a uniform rectangular array, multi-user AD MIMO-NOMA with 64 antennas has the same sum-rate performance as AD-DBF with 192 antennas. Accordingly, AD MIMO-NOMA can be an alternative solution to adding massive antennas and offers affordable hardware complexity and power consumption since DBF requires as many radio-frequency (RF) chains as antennas. Third, we shift our focus to using LSAAs with hundreds of antennas at the BS (e.g., 128, 256) to increase the BF gain, thus the network capacity and employing hybrid BF (HBF) with limited RF chains to reduce the power consumption and the complexity. The conventional AD MIMO-NOMA - also denoted as single-beam NOMA (SB-NOMA) - generates a single beam in the direction of the NOMA cluster and thereby with LSAAs can only serve the users having very close direction due to narrow beamwidth. In this context, we give particular attention to restricting the number of RF chains in mmWave hybrid systems and limiting SB-NOMA from exploiting multi-user diversity with LSAAs. And, we propose different schemes offering additional degrees of freedom (DoFs) for SB-NOMA in mmWave hybrid systems with only the knowledge of angular information. The former adopts time-division multiple access (TDMA) and exploits the time-domain resources by serving the clusters with high spatial interference within different time-slots. However, the use of TDMA with SB-NOMA imposes more signal processing and synchronization complexity at the transmitter and the receiver sides. To tackle this issue, we leverage the multi-beam (MB)-NOMA concept to provide DoFs using only NOMA and design joint SB- and MB-NOMA schemes to benefit from NOMA multi-user diversity, regardless of the cell load and users location. In particular, we propose two different schemes: the former gives a significant sum-rate performance with respect to DBF, while the latter reduces the number of active RF chains, i.e., power consumption. Simulation results show that the proposed TDMA-based scheme achieves a sum-rate gain of up to 83% over the existing TDMA-based one. Both proposed schemes offer DoFs to SB-NOMA in mmWave hybrid systems and achieve good spectral and energy efficiencies.

Keywords: Massive MIMO, mmWave communications, non-orthogonal multiple access, precoding, beamforming, partial channel state information.

Contents

Contents	VI
List of Figures	XI
List of Tables	XV
List of Abbreviations	XVII
Notations	XX
1 Introduction	1
1.1 Next-generation wireless networks	1
1.1.1 MmWave communications	2
1.1.2 Massive MIMO	2
1.1.3 Non-orthogonal multiple access	2
1.1.4 Massive MIMO-NOMA-aided mmWave communications	2
1.1.5 Challenges in mmWave massive MIMO-NOMA systems	3
1.2 Thesis contributions	3
1.3 List of publications	5
1.3.1 Journal publication	5
1.3.2 Conference publications	5
1.4 Organization of the thesis	5
2 Background	7
2.1 Introduction	7
2.2 Uniformly excited phased arrays	7
2.2.1 Array steering vector	8
2.2.2 Array factor and radiation pattern	9
2.2.3 Beamwidth	12
2.3 Key aspects of massive MIMO systems	13
2.3.1 Enhancing the spectral efficiency	14
2.3.2 Benefits of multiple antennas in wireless communications	14
2.3.3 Multi-user MIMO beamforming	16

2.3.4	From conventional to massive MIMO	19
2.4	Non-orthogonal multiple access	19
2.5	MmWave channels	22
2.5.1	Review of sub-6 GHz channel models and simulators	23
2.5.2	Review of mmWave channel models and simulators	24
2.5.3	NYUSIM mmWave channel model and simulator	24
2.5.3.1	Time-cluster and spatial-lobe approach	25
2.5.3.2	Input parameters	26
2.5.3.3	Output parameters	27
2.6	Summary	28
3	Angle-Domain Digital Beamforming for MmWave Massive MIMO Systems	29
3.1	Introduction	29
3.2	Channel and system model	31
3.2.1	Channel model	32
3.2.2	Data transmission model	32
3.2.3	Fully digital beamforming	32
3.2.3.1	Conjugate beamforming	33
3.2.3.2	Angle-domain digital beamforming	33
3.3	Sum-rate in mono- and multi-path environments	34
3.3.1	Mono-path environment	34
3.3.2	Multi-path environment	35
3.4	Illustrative results and discussions	36
3.4.1	Impact of the number of users	37
3.4.1.1	Mono-path environment	37
3.4.1.2	Multi-path environment	39
3.4.2	Impact of the number of transmit antennas	42
3.4.3	Impact of the angle spread	44
3.5	Summary	45
4	Multi-User Angle-Domain MIMO-NOMA for mmWave Communications	46
4.1	Introduction	46
4.1.1	Related works	47
4.1.2	Contributions	48
4.2	System description and problem statement	50
4.2.1	Channel Model	51
4.2.2	Data transmission	51
4.2.3	AD MIMO-NOMA with SIC	52
4.3	ADI-based performance metrics	54
4.3.1	ADI-based channel quality metric	54
4.3.2	ADI-based sum-throughput metric	55

4.3.3	ADI-based inter-user spatial interference metric	56
4.4	User clustering algorithm	58
4.4.1	2-user clustering algorithm	59
4.4.2	Multi-user clustering algorithm	59
4.4.3	Computational complexity	60
4.5	Multi-user power allocation technique	61
4.5.1	Problem formulation	61
4.5.2	Fixed inter-cluster power allocation	64
4.5.3	FCSI- and ADI-based intra-cluster power allocation	64
4.5.4	Performance comparison: FCSI- versus ADI-based schemes	66
4.5.4.1	Mono-path environment	67
4.5.4.2	Multi-path environment	67
4.6	Illustrative results and discussions	67
4.6.1	Channel estimation overhead	67
4.6.2	2D AD MIMO-NOMA performance	68
4.6.2.1	Performance of user ordering strategy	68
4.6.2.2	Performance of power allocation technique	69
4.6.2.3	Average spectral efficiency versus total number of users	71
4.6.2.4	Number of users per NOMA cluster	73
4.6.2.5	Average spectral efficiency versus spatial interference threshold	74
4.6.3	3D AD MIMO-NOMA performance	75
4.6.3.1	Average spectral and energy efficiencies	75
4.6.4	AD MIMO-NOMA complexity	77
4.7	Summary	78
5	Angle-Domain Hybrid Beamforming-based MmWave Massive MIMO-NOMA	80
5.1	Introduction	80
5.1.1	Related works	81
5.1.2	Contributions	83
5.2	System model	84
5.2.1	Hybrid beamforming design	85
5.2.2	HBF-based MIMO-NOMA system model	85
5.2.3	Problem statement and proposed solutions	86
5.3	Joint β -based SB-NOMA and TDMA scheme	87
5.4	Joint angle-domain SB- and MB-NOMA schemes	89
5.4.1	Multiple analog beams using beam splitting	89
5.4.1.1	Beam splitting with ULA	89
5.4.1.2	Beam splitting with URA	90
5.4.1.3	Illustrative representations	91
5.4.2	2-user scenario: SB-NOMA Vs MB-NOMA Vs SDMA	91
5.4.2.1	System sum-rate	92

5.4.2.2	Performance analysis	94
5.4.3	System model of the multi-user joint SB- and MB-NOMA scheme	97
5.4.3.1	SB-MB-NOMA scheme	98
5.4.3.2	SB-MB-NOMA-RRFC scheme	98
5.4.4	Low-complex MB-NOMA group clustering	99
5.5	Illustrative results and discussions	100
5.5.1	2D HBF-based MIMO-NOMA performance	101
5.5.2	3D HBF-based MIMO-NOMA performance	104
5.5.2.1	Impact of the antenna array architecture	105
5.5.2.2	Impact of the antennas number	107
5.6	Summary	108
6	Conclusion and Perspectives	110
6.1	Thesis Summary	110
6.2	Future Research Perspectives	112
6.2.1	Short-term studies:	112
6.2.2	Long-term studies:	112
	Appendix A	114
	Bibliography	115

List of Figures

1	Système MIMO-NOMA dans le domaine angulaire.	XXVI
2	Efficacité spectrale et énergétique en fonction du nombre d'utilisateurs, respectivement, pour 2-user and multi-user AD MIMO-NOMA utilisant URA avec $M_x = 32$ et $M_z = 2$ and for C-AD-DBF utilisant URA avec $M_x = 32$ et $M_z = \{2, 4, 5, 6\}$. Les deux figures utilisent la même légende présentée dans la figure Fig. 2a. . . .	XXVII
3	SB- and MB MIMO-NOMA conjoint.	XXVIII
2.1	Representation of the coordinate system and the phased planar array located in the xoz plane.	8
2.2	Representation of the inter-element path difference along x-, y-, and z-axis experienced by the plane wave upcoming from the $\vec{\Theta}$ angle w.r.t the antenna array. .	9
2.3	Illustration of the radiation pattern resulting from the product of the individual antenna pattern and the array factor of an array with four directive antennas. The figure is taken from [24].	11
2.4	Omnidirectional antenna propagation in a free space [31]	12
2.5	Illustration of $3dB$ -beamwidth	12
2.6	Normalized pattern of a ULA array having $M \in \{16, 32, 64\}$ antenna elements and generating a beam at the azimuth angle $\theta_0 \in \{30^\circ, 60^\circ\}$	13
2.7	Normalized 3D array pattern of a beam scanned at $\vec{\Theta}_0 = (60^\circ, 30^\circ)$ with (a) 32×4 URA and (c) 32×6 URA.	13
2.8	Illustration of (a) SU-MIMO and (b) MU-MIMO systems in downlink [24]. . . .	15
2.9	Downlink MU-MIMO system.	17
2.10	Block diagram of the (a) RF analog beamforming, (b) baseband digital beamforming and (c) hybrid RF-baseband beamforming technique.	17
2.11	RF precoding for hybrid (a) fully-connected and (b) sub-connected structures. . .	18
2.12	Comparison between OMA and NOMA scheduling.	20
2.13	A 3-user SISO-NOMA downlink transmission with SIC decoding.	22
2.14	Representation of multi-path environment.	23
2.15	Graphical user interface of NYUSIM Version 2.	25
2.16	Omnidirectional power delay profile (PDP).	26
2.17	AoD and AoA power spectrum.	26
3.1	Downlink MU-MIMO system.	31

3.2	System model of downlink MU-MIMO digital beamforming.	33
3.3	Average sum-rate achieved by a mmWave m-MU-MIMO system using either DBS or CB in an urban mono-path environment. The BS adopts a ULA or a URA array with $M = 64$ antennas.	38
3.4	Azimuth representation of the spatial distribution of UEs (*), on the xoy plane, with their ULA beam patterns, using DBS for the two realizations in a mono-path environment. Each UE is represented by a colored * and its corresponding beam has the same color. Note that the scale corresponds to the separation distance between UEs and BS in meter, where BS is located at the center of the circle. . .	39
3.5	Average sum-rate achieved by a mmWave m-MU-MIMO system using either DBS or CB in an urban multi-path environment. The BS adopts a ULA or a URA array with $M = 64$ antennas.	40
3.6	Azimuth representation of the spatial distribution of paths, on the xoy plane, for one realization in an urban multi-path environment. All paths belonging to the same UE channel are identified by the same color. The used scale corresponds to their channel gain.	41
3.7	Azimuth representation of the spatial distribution of UEs (*), on the xoy plane, with their ULA beam patterns, using DBS, for one realization. Each UE is illustrated by a colored * and its corresponding beam has the same color. . . .	42
3.8	Average sum-rate achieved using either DBS or CB against the number of antennas, M , in a rural or urban multi-path environment when the BS is equipped with ULA and serves $K = 32$ users.	43
3.9	Average sum-rate achieved using either DBS or CB against the number of vertical antennas, M_z , in a rural or urban multi-path environment when the BS is equipped with URA and serves $K = 32$ users.	43
3.10	Impact of the angle spread on the average sum-rate in a two-path channel when a BS equipped with 64 linear antennas along the x -axis serves $K \in \{16, 32\}$. The sum-rate is averaged over 5000 channel realizations.	44
4.1	Multi-user MIMO-NOMA system.	50
4.2	Illustration of the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ of the beam pointed toward UE u and the value $AF_{(\theta_{1,u})}(\theta_{1,k})$ for the spatial direction $\theta_{1,k}$ corresponding to $\beta_{k,u}$ of UE k and UE u	57
4.3	Illustration of the normalized array factor $ AF_{(\theta_{1,u})}^n(\theta) $ of the beam generated toward UE u for $\theta_{1,u} \in \{50^\circ, 110^\circ\}$, with $M = 32$	58
4.4	Average system spectral efficiency of 2-user AD MIMO-NOMA for different ordering strategies using distance, effective channel gain $ \mathbf{h}_{l,c}\mathbf{a}_c ^2$, ADI-based or FCSI-based channel quality with FPA ($\psi = 0.4$).	69
4.5	Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with FCSI NOMA.	70

4.6	Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with ADI NOMA.	71
4.7	Average system spectral efficiency versus the total number of users, K , for $\beta_0 = 0.5$	72
4.8	Average number of clusters (a), cluster's percentage distribution per type for 2-user (b) and multi-user AD MIMO-NOMA (c) versus the total number of users, K , for $\beta_0 = 0.5$	74
4.9	Average system spectral efficiency of AD MIMO-NOMA with ADI NOMA, for 2-user and multi-user clusters, when $K \in \{10, 20, 40\}$ users exist.	75
4.10	Average system spectral and energy efficiencies for 2-user and multi-user AD MIMO-NOMA using URA with $M_x = 32$ and $M_z = 2$ and for C-AD-DBF using URA with $M_x = 32$ and $M_z = \{2, 4, 5, 6\}$ in a rural environment. Both figures use the same legend presented in Fig. 4.10a.	77
5.1	HBFB-based SB-NOMA-MIMO scheme.	84
5.2	Vertically array splitting technique with URA.	90
5.3	Antenna array pattern response for multiple beam pointed at users k and k' , and for a single beam pointed toward user k'' with (a) ULA and (b) URA.	91
5.4	SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\Delta\theta$ and $\beta_{1,2}$ with 128×1 ULA.	95
5.5	SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\Delta\theta$ and $\Delta\phi$ with 16×16 URA.	95
5.6	SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\beta_{1,2}$ with 16×16 URA.	96
5.7	Joint SB- and MB MIMO-NOMA scheme.	98
5.8	Flowchart of the proposed SB-MB-NOMA and SB-MB-NOMA-RRFC schemes.	99
5.9	Three different 2-by-2 selection strategies.	100
5.10	The sum-rate performance of different MB-NOMA UC algorithms versus the number of users for the SB-MB-NOMA-RRFC scheme. The sum-rate is averaged over 5,000 trials.	100
5.11	(a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users with a 128×1 ULA array and 20 RF chains at the BS.	102
5.12	(a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users of the different schemes with 16×8 URA and 32×4 URA and 20 RF chains at the BS.	105
5.13	(a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users with 32×8 URA and 20 RF chains at the BS.	106
5.14	(a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of horizontal antennas, M_x , when $K = 40$ users exist in the cell with $M_z = 8$ vertical antennas and $N_{RF} \in \{20, 25\}$ RF chains at the BS.	107

- 5.15 (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of vertical antennas M_z when $K = 40$ users exist in the cell with $M_x = 32$ horizontal antennas and $N_{RF} \in \{20, 25\}$ RF chains at the BS.108

List of Tables

3.1	Simulation parameters	37
3.2	System sum-rate, rate and channel gain of each UE for two different realizations where five UEs are served simultaneously with ULA in a mono-path environment.	38
3.3	Sum-rate and rate of each UE where seven UEs are served simultaneously by ULA or 16×4 URA and using either DBS or CB. These obtained values correspond to for the same realization	40
4.1	Simulation parameters	67
4.2	Channel Estimation Overhead. Note that $N_e = 1[\theta_{1,k}]$, (resp. $2[(\theta_{1,k}, \phi_{1,k})]$) with ULA (resp. URA).	68
4.3	The complexity of 2-user and multi-user β -UC algorithms and the number of NOMA clusters of both schemes, when the BS serves $K \in \{10, 20, 30, 40\}$ users. .	78
4.4	Relative percentage of NOMA users and SIC users for the multi-user MIMO-NOMA scheme, as compared to the 2-user scheme, when the BS serves $K \in \{20, 30, 40\}$ users.	78
5.1	Simulation parameters	101

List of Abbreviations

1G	First generation
2D	Two dimensional
3D	Three dimensional
3GPP	Third generation partnership project
4G	Fourth generation
5G	Fifth generation
ABF	Analog beamforming
AD	Angle-domain
AD-DBF	AD-digital beamforming
ADI	AD information
AoA	Angle-of-arrival
AoD	Angle-of-departure
AWGN	Additive white Gaussian noise
BF	Beamforming
bps	Bit-per-second
BS	Base station
CB	Conjugate beamforming
CIR	Channel impulse response
CSI	Channel state information
CSIT	Channel state information at the transmitter
dB	Decibel
DBF	Digital beamforming
DBS	Digital beamsteering
DL	Downlink
EE	Energy efficiency
eMBB	Enhanced Mobile Broadband

FCSI	Fixed CSI
FDD	Frequency division duplex
FDMA	Frequency division multiple access
FD-MIMO	Full-dimensional MIMO
FPA	Fixed power allocation
FTPA	Fractional transmit power allocation
ICT	Information and Communications Technology
IoT	Internet-of-things
HBF	Hybrid beamforming
KKT	Karush-Kuhn-Tucker
LoS	Line-of-sight
LSAA	large-scale antenna array
LTE	Long-term evolution
mMTC	Massive machine-type communications
MIMO	Multiple-input multiple-output
m-MIMO	massive MIMO
mmWave	Millimeter-wave
MPC	Multi-path component
MU-MIMO	Multi-user MIMO
NLoS	Non-LoS
NYUSIM	New York University simulator
NOMA	Non-orthogonal multiple access
OFDM	Orthogonal frequency division multiplexing
OMA	Orthogonal multiple access
PA	Power allocation
PAPR	Peak-to-average power ratio
PD-NOMA	Power-domain NOMA
QoS	Quality of service
RF	Radio-frequency

RMa	Rural macrocell
RMS	Root-Mean-Square
Rx	Receiver
SCM	Superposition coded modulation
SCSI	Statistical CSI
SE	Spectral efficiency
SINR	Signal-to-interference-plus-noise ratio
SLNR	Signal-to-leakage-plus-noise ratio
SNR	Signal-to-noise ratio
SIC	Successive interference cancellation
SVD	Single value decomposition
SWIPT	Simultaneous wireless information and power transfer
TDD	Time division duplex
TDMA	Time division multiple access
Tx	Transmitter
UC	User clustering
UE	User equipment
UL	Uplink
ULA	Uniform linear array
UMi	Urban microcell
UMa	Urban macrocell
URA	Uniform rectangular array
URLLC	Ultra-reliable low-latency communications

Mathematical Notations

x	Scalar
$\mathbf{x}$	Vector
$\mathbf{X}$	Matrix
$\mathbf{X}^{-1}$	Inverse of matrix $\mathbf{X}$
$\mathbf{X}^*$	Complex conjugate of $\mathbf{X}$
$\mathbf{X}^T$	Transpose of $\mathbf{X}$
$\mathbf{X}^H$	Conjugate transpose of $\mathbf{X}$
$Tr(\mathbf{X})$	Trace of $\mathbf{X}$
$ x $	Amplitude of complex number x
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbb{C}$	Complex field
$[\cdot]$	Integer part
$\ \cdot\ $	Euclidean norm
$\mathbf{card}(\cdot)$	Cardinality of a set
$\otimes$	Kronecker product
$\mathcal{N}(0, \sigma_n^2)$	zero-mean white Gaussian distribution with variance σ^2
$\mathcal{U}\{a, b\}$	Discrete uniform distribution of the integers between a and b
e^x	Exponential function of x
$\log_x(\cdot)$	Logarithmic function of base x
$\mathcal{O}(\cdot)$	Complexity order

Résumé Long

Les réseaux de communications sans fils sont devenus une infrastructure sociale essentielle, moteur de l'économie numérique et marquant la vie quotidienne des gens. Cette tendance exige du réseau cellulaire qu'il relie les industries à la communauté des technologies de l'information et des communications et qu'il améliore l'expérience d'utilisateur [1]. De plus, l'Internet des objets (IoT) est apparu comme un nouveau paradigme, de sorte qu'un nombre massif des appareils et des capteurs sont et seront connectés au réseau. En d'autres termes, le système sans fil est passé de la connexion des personnes à la connexion des personnes et des objets. En outre, divers cas d'utilisation émergents ont été introduits, tels que les services de santé en ligne, les applications de réalité augmentée/virtuelle, l'informatique en nuage, les robots télécommandés, la conduite automobile, les capteurs de surveillance et de contrôle, etc. Pour répondre à la croissance rapide du trafic de données, à l'explosion du nombre de dispositifs connectés et aux exigences de qualité de service, une nouvelle génération de réseau sans fil est appelée à développer et à adopter de nouvelles technologies efficaces. Dans ce contexte, la cinquième génération (5G) de communications cellulaires et au-delà a émergé rapidement pour répondre à ces exigences strictes, en proposant des technologies prometteuses qui font de la 5G une révolution des générations précédentes plutôt qu'une évolution [4, 5]. En particulier, le futur réseau sans fil adopte les bandes d'ondes millimétriques (mmWave), le système massive multiple-input multiple output (MIMO), de nouvelles techniques d'accès multiple comme l'accès multiple non orthogonal (NOMA), le déploiement de cellules denses, etc. L'intégration des trois premières technologies, à savoir "systèmes MIMO-NOMA massifs à ondes millimétriques", a fait l'objet d'une attention particulière dans la littérature récente afin d'améliorer de manière significative la capacité des réseaux et de fournir une qualité de service pour les futurs services. Le thème général de cette thèse tourne autour de ce nouveau type de communication.

Objectifs de la thèse et contributions

Cette thèse vise à relever les défis critiques des systèmes MIMO-NOMA massifs à ondes millimétriques en transmission descendante qui résultent de l'utilisation d'un réseau d'antennes à grande échelle, à savoir la complexité de la conception du précodage, les contraintes matérielles et le surcoût de l'estimation du canal. Tout au long de cette thèse, nous exploitons les principales caractéristiques des canaux à ondes millimétriques, c'est-à-dire la directionnalité et le taux de blocage élevé. dans ce contexte, l'information géométrique caractérisé par l'angle du trajet

de la transmission principale est une information de l'état du canal partiel prometteur, dont la surcharge de rétroaction n'augmente pas avec le nombre d'antennes. De plus, la périodicité d'estimation d'un canal de transmission (en fréquence ou en temps) dépend intrinsèquement de la fréquence. En revanche, la mise à jour de l'estimation des directions ne dépend que de la mobilité de chaque utilisateur. L'objectif principal de cette thèse est de prouver l'intérêt d'utiliser uniquement l'information angulaire relative aux positions des utilisateurs pour réaliser des méthodes de précodage, de regroupement et d'ordonnancement des utilisateurs et d'allocation de puissance peu complexes pour des systèmes MIMO-NOMA massifs en ondes millimétriques avec des techniques de formation du faisceau numérique (digital beamforming (DBF) en anglais) ou de BF hybride (hybrid beamforming (HBF) en anglais). De plus, nous avons appliqué des canaux d'ondes millimétriques 2D et 3D à l'aide du simulateur réaliste et statistique de canaux d'ondes millimétriques, appelé NYUSIM [22, 23], développé par l'Université de New York.

Les principales contributions de cette thèse peuvent être résumées comme suit

- Tout d'abord, nous considérons les systèmes MIMO massifs à ondes millimétriques avec des techniques DBF. Afin de réduire la complexité de la conception du précodage et le surcoût lié à l'acquisition des canaux, nous analysons le comportement spatial et la potentialité de la technique AD-DBF (angle-domain DBF en anglais) dans des environnements mono- et multi-trajets. Plus précisément, la technique AD-DBF n'utilise que des déphaseurs numériques et ne nécessite que les informations angulaires des utilisateurs. Comme référence, nous considérons la technique conjugate beamforming (CB) qui maximise le gain de beamforming au niveau de l'utilisateur souhaité, sur la base de la connaissance complète du canal de transmission. Contrairement aux travaux antérieurs [24] qui adoptent un modèle de canal simplifié pour les ondes millimétriques, nous utilisons le modèle et le simulateur de canal réaliste et statistique NYUSIM.
- Deuxièmement, nous considérons la limitation de la technique AD-DBF, en particulier dans une cellule avec beaucoup des utilisateurs, c'est-à-dire avec une interférence spatiale entre utilisateurs voisins élevée, lorsque des dizaines d'antennes sont mises en œuvre au niveau de l'émetteur. Au lieu d'augmenter le nombre d'antennes d'émission, et donc la complexité et le coût du matériel, nous proposons une solution basée sur NOMA qui exploite le domaine de la puissance et dessert plusieurs utilisateurs dans le même domaine spatial. En particulier, nous tirons parti des caractéristiques des ondes millimétriques et nous concevons un nouveau schéma AD MIMO-NOMA qui peut gérer une connectivité massive dans un scénario encombré. Dans la deuxième étude, nos contributions peuvent être résumées comme suit
 - Nous définissons d'abord de nouvelles mesures basées sur l'angle, qui tirent parti des principales caractéristiques des canaux à ondes millimétriques: la qualité du canal de l'utilisateur, le débit total et l'interférence spatiale entre utilisateurs, uniquement sur la base de l'information angulaire.
 - Nous concevons une stratégie d'ordonnancement des utilisateurs dans le domaine an-

gulaire pour le protocole NOMA. Contrairement aux stratégies basées sur le CSI partiel existant dans la littérature, notre schéma prend en compte l'interférence spatiale inter-clusters.

- Nous développons des algorithmes de regroupement à deux utilisateurs et à plusieurs utilisateurs pour réduire l'interférence inter-utilisateurs. À notre connaissance, les algorithmes que nous proposons sont les premiers algorithmes de regroupement d'utilisateurs qui utilisent l'interférence spatiale inter-utilisateurs basée sur l'angle. Cette dernière ne nécessite que des informations angulaires et inclut la distance angulaire et la largeur de faisceau du lobe principal dans une métrique unique.
- Nous utilisons le débit total qui est dérivé en se basant sur l'angle pour concevoir une technique d'allocation de puissance sous-optimale qui est cohérente avec le protocole NOMA dans le domaine de la puissance et garantit une annulation d'interférences efficace.
- Troisièmement, nous profitons de la technologie massive MIMO pour mettre en œuvre des réseaux à grande échelle (par exemple, avec 128 ou 256 antennes) et augmenter le gain de la BF et donc la capacité du réseau. Cependant, DBF nécessite autant de chaînes radio-fréquence (RF) que d'antennes, ce qui entraîne une consommation d'énergie élevée ainsi qu'une complexité et un coût matériel inabordables. C'est pourquoi, dans cette étude, nous passons au HBF avec un nombre de chaînes RF limité. Nous étudions le MIMO-NOMA massif basé sur HBF dans le domaine angulaire pour les communications en ondes millimétriques. Le MIMO-NOMA conventionnel - également appelé NOMA à faisceau unique (single-beam NOMA (SB-NOMA)) - génère un faisceau unique dans la direction du cluster NOMA. Avec des antennes massives, la largeur de faisceau est très étroite et, par conséquent, en utilisant le schéma SB-NOMA, seuls les utilisateurs ayant des angles similaires peuvent être servis simultanément dans la même groupe de NOMA. Par conséquent, en raison d'un nombre de chaînes RF limité dans les systèmes hybrides à ondes millimétriques, SB-NOMA avec des antennes massives ne peut pas gérer la connectivité de tous les utilisateurs, en particulier dans une cellule encombrée. Pour résoudre ce problème, nous proposons différents schémas qui offrent des degrés de liberté supplémentaires pour SB-NOMA dans les systèmes hybrides à ondes millimétriques avec seulement la connaissance de l'information angulaire. Dans la troisième étude, nos contributions peuvent être résumées comme suit
 - Nous concevons d'abord un schéma conjoint SB-NOMA et TDMA (Time-division multiple access) exploitant les ressources du domaine temporel pour assurer la connectivité de tous les utilisateurs même lorsque le nombre total de groupes (mono-utilisateur et SB-NOMA) est supérieur à celui des chaînes RF. Contrairement aux travaux précédents, le schéma que nous proposons peut s'appliquer à n'importe quelle architecture de réseau d'antennes.
 - Nous tirons ensuite parti du concept de NOMA multifaisceaux (multi-beam NOMA

(MB-NOMA)) pour réduire la complexité du traitement du signal lors de l'application de deux techniques d'accès multiple différentes. Plus précisément, nous proposons deux schémas conjoints SB- et MB-NOMA pour bénéficier de la diversité multi-utilisateurs NOMA, indépendamment de la charge de la cellule et de l'emplacement de l'utilisateur. Le premier donne une amélioration de performance significative en débit par rapport à AD-DBF, tandis que le second réduit le nombre de chaînes RF actives, c'est-à-dire la consommation d'énergie.

Structure du manuscrit et résumé des chapitres

Le manuscrit de thèse est composé de six chapitres, qui sont résumés dans les paragraphes suivants

Chapitre 1 - Introduction Générale

Le chapitre 1 est un chapitre introductif, qui positionne le travail de thèse dans le contexte des réseaux sans fil de nouvelle génération. Nous passons en revue les technologies clés: communication à ondes millimétriques, système massive MIMO et NOMA. Les contributions de la thèse sont également résumées, et la liste des publications est fournie. Dans la dernière partie du chapitre, l'organisation de la thèse est présentée.

Chapitre 2 - Contexte

Ce chapitre couvre différents sujets en lien avec le travail présenté dans cette thèse. Tout d'abord, nous donnons les principales caractéristiques des réseaux linéaires et planaires équidistants avec une excitation uniforme. Puis, nous présentons les principaux aspects des systèmes MIMO conventionnels et massifs pour les transmissions en liaison descendante. Nous classons les différents types de techniques de précodage MIMO multi-utilisateurs. Ensuite, nous présentons les principes fondamentaux de la transmission NOMA dans le domaine de la puissance sur la liaison descendante. Enfin, un examen approfondi des modèles de canaux sous-6 GHz et à ondes millimétriques est fourni. En particulier, le simulateur de modèle de canal mmWave réaliste et statistique, appelé NYUSIM, sera pris en compte tout au long de cette thèse. Enfin, une explication détaillée des différentes étapes pour générer le modèle de canal est présentée.

Chapitre 3 - Angle-Domain Digital Beamforming pour les systèmes MIMO massifs à ondes millimétriques

Dans ce chapitre, nous nous concentrons sur la réduction de la complexité du précodage et de la surcharge de l'estimation du canal en exploitant la potentialité de l'information angulaire dans les canaux à ondes millimétriques hautement directionnels. Par conséquent, tout au long

de ce chapitre, nous adoptons AD-DBF [92, 93], le formateur du faisceau un précodeur à enveloppe constante qui n'utilise que des déphaseurs numériques et ne nécessite qu'une estimation géométrique partielle du canal, principalement l'information angulaire des utilisateurs par rapport à la station de base. Contrairement aux travaux antérieurs qui utilisaient un modèle de canal simplifié pour les ondes millimétriques ([92, 93]), nous considérons le modèle de canal réaliste et statistique NYUSIM de l'Université de New York. NYUSIM fournit des modèles de canaux MU-MIMO 2D et 3D utilisant différentes architectures de réseaux d'antennes linéaires et rectangulaires. Comme référence, nous considérons le précodeur CB, qui est moins complexe que ZF et RZF et maximise le gain BF à l'utilisateur désiré, sur la base d'une connaissance complète du canal; [50, 51]. L'objectif de ce chapitre est d'étudier le comportement spatial de AD-DBF et sa potentialité pour les systèmes m-MIMO en ondes millimétriques, ainsi que de fournir une référence pour le chapitre suivant.

En s'appuyant sur des simulations basées sur des aspects géométriques, nous avons constaté que la technique AD-DBF est une technique de beamforming prometteuse qui permet d'équilibrer les performances, la complexité matérielle et la surcharge d'estimation des canaux. Dans les environnements de propagation à trajet unique, AD-DBF donne un débit modéré pour chaque utilisateur, tandis que CB renforce les utilisateurs forts et marginalise les faibles. Dans les environnements à trajets multiples, pour AD-DBF, même si les utilisateurs sont bien séparés dans l'espace, il y a des interférences résultant des trajets non directs (non-line-of sight paths (NLoS)). Cependant, CB maximise le gain de beamforming en exploitant la diversité des chemins NLoS, alors que AD-DBF n'exploite pas leur potentiel et perd une partie de l'énergie rayonnée. Avec un grand nombre d'antennes par rapport à celui des utilisateurs, AD-DBF surpasse CB même dans des environnements à trajets multiples. De plus, AD-DBF présente une performance significative en termes de débit total dans les environnements hautement directionnels, où tous les chemins partent de directions proches. Cependant, dans la pratique, en raison de la complexité et du coût élevés du matériel, le nombre d'antennes ne peut pas approcher l'infini, de sorte que la station de base ne peut pas éliminer complètement les interférences entre les utilisateurs. Avec un nombre fini d'antennes, surtout dans une cellule encombrée, les performances de AD-DBF se dégradent en raison de l'interférence inter-utilisateur élevée causée par les faisceaux voisins. Pour résoudre ce problème, nous proposons, dans le chapitre suivant, un schéma de gestion des interférences sans déploiement d'antennes supplémentaires.

Chapitre 4 - MIMO-NOMA dans le domaine angulaire pour les communications en ondes millimétriques

Dans ce chapitre, nous visons à concevoir un schéma AD MIMO-NOMA pour les systèmes multi-utilisateurs à ondes millimétriques afin de réduire les interférences entre utilisateurs et d'améliorer les performances AD-DBF sans ajouter d'antennes supplémentaires. En particulier, nous exploitons les principales caractéristiques des canaux à ondes millimétriques et nous proposons de nouvelles mesures basées sur l'angle pour concevoir et développer des techniques

efficaces de regroupement et d'ordonnancement des utilisateurs et d'allocation de puissance. Ces techniques ne nécessitent que des informations angulaires et peuvent être utilisées pour les architectures d'antennes linéaires et rectangulaires. La figure 1 illustre notre schéma MIMO-NOMA proposé: grâce à la définition d'une métrique d'interférence spatiale inter-utilisateurs, les utilisateurs sont regroupés dans des groupes, adressés simultanément par les faisceaux de AD-DBF. Le NOMA est utilisé dans chaque groupe pour gérer l'interférence et l'accès multiple.

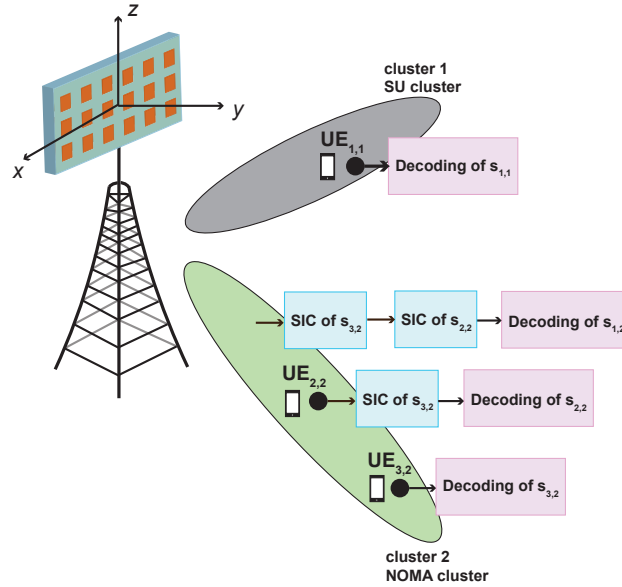


Figure 1: Système MIMO-NOMA dans le domaine angulaire.

Nous avons d'abord exploité les principales caractéristiques des canaux à ondes millimétriques, à savoir la forte directionnalité et le blocage potentiel. Ensuite, nous avons dérivé un ensemble de mesures de performance basées sur l'angle, telles que l'interférence spatiale inter-utilisateur, la qualité du canal utilisateur et le débit total. Ces mesures ont ensuite participé à la conception des techniques de regroupement et d'ordonnancement des utilisateurs et d'allocation de puissance. Plus précisément, l'algorithme de regroupement des utilisateurs regroupe les utilisateurs présentant une forte interférence spatiale. Et pour lesquels la technique NOMA va s'appliquer. Une nouvelle stratégie d'ordonnancement des utilisateurs est proposée en utilisant la métrique de qualité du canal basée sur l'angle. En outre, nous avons conçu une méthode d'allocation de puissance qui maximise le débit total basé sur l'angle. Dans ce chapitre, nous avons étudié non seulement le cas des groupes NOMA à deux utilisateurs, mais aussi celui des groupes NOMA à utilisateurs multiples. Des résultats numériques détaillés ont montré que le schéma AD MIMO-NOMA multi-utilisateurs proposé améliore considérablement les performances du système MU-MIMO en ondes millimétriques en obtenant une augmentation de 39% de débit total lorsque le nombre d'utilisateurs est proche de celui des antennes. En outre, nous avons constaté que la stratégie d'ordonnancement des utilisateurs proposée surpasse les autres stratégies qui utilisent un csi partiel, et que l'allocation de puissance basée sur l'angle permet une décodage SIC efficace. Enfin, les résultats ont montré le potentiel du NOMA par rapport

à OMA pour exploiter la diversité multi-utilisateurs dans une cellule encombrée. En outre, AD MIMO-NOMA est une solution alternative à l'ajout d'antennes supplémentaires.

Exemples de Résultats

Les figures Fig. 2a et Fig. 2b représentent l'efficacité spectrale et énergétique en fonction du nombre d'utilisateurs, respectivement, pour AD MIMO-NOMA utilisant un réseau d'antennes rectangulaire avec deux antennes en vertical, et pour AD-DBF utilisant un nombre différent d'antennes en vertical. Notez que les différents schémas AD MIMO-NOMA et AD-DBF ont le même nombre d'antennes en horizontal, c'est-à-dire $M_x = 32$ et un nombre différent d'antennes en vertical, M_z .

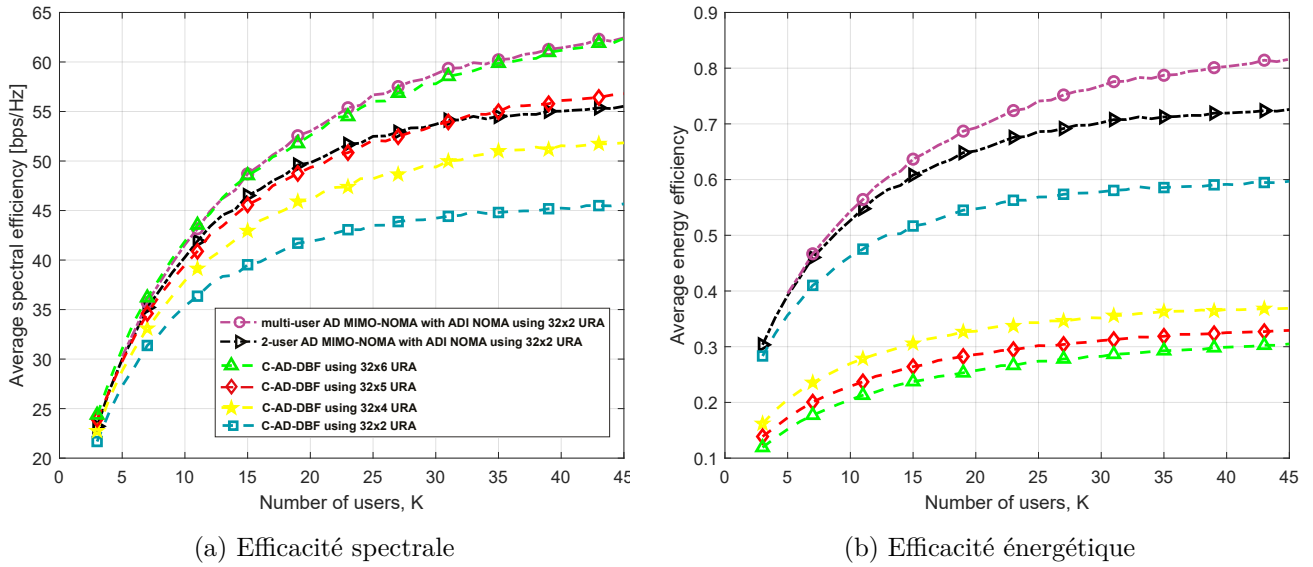


Figure 2: Efficacité spectrale et énergétique en fonction du nombre d'utilisateurs, respectivement, pour 2-user and multi-user AD MIMO-NOMA utilisant URA avec $M_x = 32$ et $M_z = 2$ and for C-AD-DBF utilisant URA avec $M_x = 32$ et $M_z = \{2, 4, 5, 6\}$. Les deux figures utilisent la même légende présentée dans la figure Fig. 2a.

En utilisant 96 (resp. 128) d'antennes en moins, le schéma AD MIMO-NOMA à 2 utilisateurs (resp. multi-utilisateurs) proposé avec $M_z = 2$ a presque les mêmes performances en efficacité spectrale que C-AD-DBF avec $M_x = 32$ et $M_z = 5$ (resp. $M_x = 32$ et $M_z = 6$), comme le montre la Fig. 4.10a. En effet, les antennes supplémentaires augmentent le gain de BF et réduisent ainsi les interférences inter-utilisateurs. En revanche, AD MIMO-NOMA peut améliorer l'efficacité spectrale sans déployer d'antennes supplémentaires en multiplexant les utilisateurs présentant un brouillage inter-utilisateur élevé dans le domaine de la puissance. Par conséquent, il est évident que le gain en efficacité énergétique du système AD MIMO-NOMA à 2 utilisateurs (resp. multi-utilisateurs) est 129% (resp. 170%) fois plus grand que celui utilisant C-AD-DBF avec $M_z = 5$ (resp. $M_z = 6$), comme le montre la Fig. 4.10b. Donc, AD MIMO-NOMA est une solution alternative pour répondre à la connectivité massive sans augmenter significativement le nombre

d'antennes.

Chapitre 5 - Massive MIMO-NOMA en ondes millimétriques basé sur hybrid beamforming dans le domaine angulaire

Dans le chapitre 5, nous avons étudié les systèmes massive MIMO-NOMA basés sur HBF avec des centaines d'antennes à la station de base (par exemple, 128, 256). Avec un réseau à multi-antennes massifs, le schéma SB-NOMA conventionnel ne peut pas exploiter la diversité multi-utilisateurs dans les systèmes hybrides à ondes millimétriques en raison de l'étroitesse des faisceaux et/ou du nombre de chaînes RF limité. Dans ce chapitre, nous avons accordé une attention particulière à ces limitations et nous avons proposé deux schémas différents pour offrir des degrés de liberté supplémentaires au schéma SB-NOMA conventionnel dans les systèmes hybrides à ondes millimétriques. Il est intéressant de noter que ces schémas ne nécessitent que la connaissance de l'information angulaire et s'appliquent aux architectures linéaires et rectangulaires, contrairement à ce qui est proposé dans la littérature. Le premier applique TDMA avec SB-NOMA et exploite les ressources du domaine temporel en desservant les groupes à forte interférence spatiale dans des créneaux temporels distincts. Pour réduire la complexité du traitement du signal à la fois à l'entrée et à la sortie lors de l'application de TDMA et du NOMA, nous avons exploité le concept du MB-NOMA et proposé des schémas conjoints SB- et MB-NOMA pour bénéficier de la diversité multi-utilisateurs du NOMA, quelle que soit la charge de la cellule et la position des utilisateurs. La figure Fig. 3 illustre le modèle du système des schémas proposés

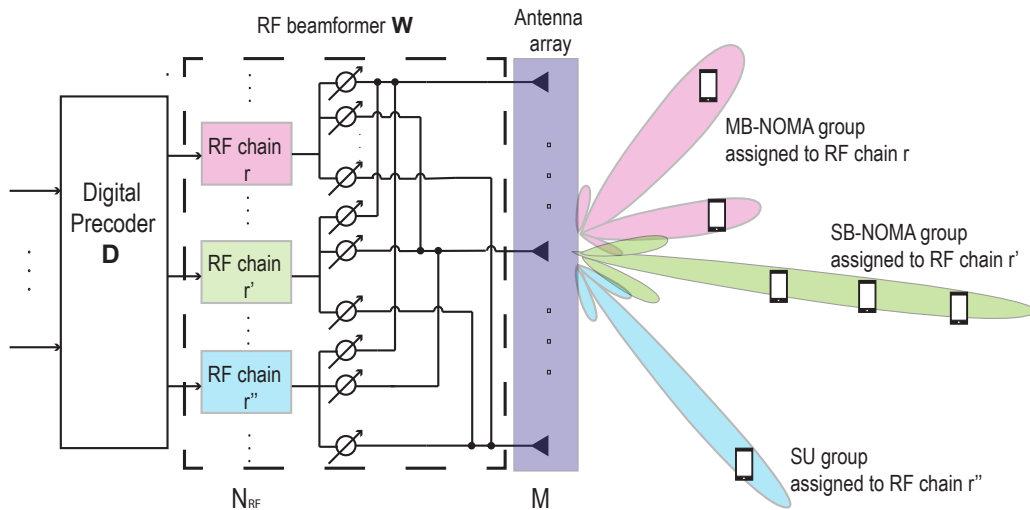


Figure 3: SB- and MB MIMO-NOMA conjoint.

Plus précisément, nous avons proposé deux schémas conjoints SB- et MB-NOMA: le premier donne une amélioration significative de débit total par rapport à DBF, tandis que le second réduit le nombre de chaînes RF actives, c'est-à-dire la consommation d'énergie. Nous avons validé les performances des schémas proposés à l'aide de simulations numériques en termes de

débit total et de nombre de chaînes RF actives en comparant à la solution proposée dans la littérature et à d'autres utilisant DBF. Les résultats de simulation ont montré que les schémas proposés permettent d'obtenir des gains de performance significatifs en termes de débit total et de nombre de chaînes RF actives, par rapport à la solution proposée dans la littérature et à d'autres schémas basés sur DBF. Par exemple, le schéma proposé, basé sur TDMA, permet d'obtenir un gain de débit total allant jusqu'à 83% par rapport au schéma basé sur TDMA existant dans la littérature. De plus, en utilisant 32×4 URA, le schéma hybride à base de faisceau unique et multiple (comme AD-HBF-SB-MB-NOMA) dépasse le schéma complètement digital (AD-DBF) de presque 23%, lorsque 40 utilisateurs sont servis dans la cellule avec 6.4 fois moins de chaînes RF actives.

Chapitre 6 - Conclusions et perspectives

Cette thèse aborde les défis critiques des systèmes massive MIMO-NOMA en ondes millimétriques, à savoir la complexité de la conception du précodage, les contraintes matérielles, la complexité de l'estimation du canal et l'overhead, qui résultent de l'utilisation d'un réseau à grande échelle. Tout d'abord, nous avons exploité la nature éparse des canaux à ondes millimétriques grâce à la haute directionnalité et au blocage potentiel, qui font de l'information angulaire un CSI partiel prometteur. Nous avons ensuite démontré l'intérêt d'utiliser uniquement cette information angulaire relative aux positions des utilisateurs pour réaliser des techniques peu complexes de précodage, de regroupement et d'ordonnancement des utilisateurs et d'allocation de puissance pour les systèmes m-MIMO-NOMA en ondes millimétriques en utilisant les techniques DBF et HBF. Pour ce faire, nous avons utilisé le simulateur réaliste et statistique de canaux de propagation sur ondes millimétriques conçu par l'université de New York NYUSIM.

Dans un premier temps, nous avons étudié le comportement spatial et la potentialité de AD-DBF dans des canaux à ondes millimétriques et avec des réseaux multi-antennes à grande échelle en utilisant NYUSIM. Nous avons constaté que AD-DBF est une technique de formation de faisceau attrayante qui offre un bon compromis entre la performance, la complexité du matériel et la surcharge d'estimation du canal. Malheureusement, ces systèmes souffrent d'une interférence inter-utilisateurs élevée lorsque les utilisateurs sont proches les uns des autres dans la direction. Deuxièmement, dans ce contexte, des schémas AD MIMO-NOMA à deux et plusieurs utilisateurs, tirant parti de NOMA dans le domaine de la puissance, sont proposés lorsque des dizaines d'antennes existent au niveau de la station de base. Tous deux visent à réduire les interférences entre utilisateurs, en particulier dans une cellule encombrée, et à améliorer l'efficacité spectrale du système. Pour ce faire, nous avons d'abord dérivé un ensemble de mesures de performance basées sur l'angle, telles que l'interférence spatiale inter-utilisateur, la qualité du canal utilisateur et le débit total. Ces mesures participent ensuite à la conception des techniques de regroupement et d'ordonnancement des utilisateurs, et d'allocation de puissance. Plus précisément, l'algorithme de regroupement des utilisateurs regroupe les utilisateurs présentant une forte interférence spatiale afin de les traiter via NOMA. Une nouvelle stratégie

d'ordonnancement des utilisateurs est proposée en utilisant la métrique de qualité du canal basée sur l'angle. En outre, nous concevons une méthode d'allocation de puissance qui maximise le débit total basé sur l'angle. Troisièmement, nous nous concentrons sur l'utilisation de réseaux avec des centaines d'antennes au niveau de la station de base (par exemple, 128 ou 256) pour augmenter le gain du BF, et donc la capacité du réseau, et sur l'utilisation de HBF avec des chaînes RF limitées pour réduire la consommation d'énergie et la complexité. SB-NOMA génère un faisceau unique dans la direction du groupe NOMA et, avec nombre massif d'antennes, il ne peut desservir que les utilisateurs ayant une direction très proche en raison de la largeur de faisceau étroite. Pour répondre à l'utilisation massive d'antennes tout en limitant le nombre de chaînes rf et en tirant parti de la diversité multi-utilisateurs nous proposons différents schémas offrant des degrés de liberté supplémentaires pour le SB-NOMA dans les systèmes hybrides à ondes millimétriques avec la seule connaissance de l'information angulaire. Le premier adopte TDMA et exploite les ressources du domaine temporel en desservant les groupes à forte interférence spatiale dans différents créneaux temporels. Cependant, l'utilisation de TDMA avec le système SB-NOMA impose une plus grande complexité de traitement du signal et de synchronisation au niveau de l'émetteur et du récepteur. Pour résoudre ce problème, nous tirons parti du concept de MB-NOMA pour fournir plus des degrés de liberté en utilisant uniquement le NOMA et nous concevons des schémas conjoints SB- et MB-NOMA pour bénéficier de la diversité multi-utilisateurs NOMA, indépendamment de la charge de la cellule et de l'emplacement des utilisateurs. En particulier, nous proposons deux schémas différents: le premier donne une performance de débit total significative par rapport au DBF, tandis que le second réduit le nombre de chaînes RF actives, c'est-à-dire la consommation d'énergie.

De nombreuses perspectives peuvent être envisagées pour compléter ce travail. L'extension de l'étude sur la voie montante est très prometteuse. En effet, la prise en compte de MIMO-NOMA dans le domaine angulaire sur la voie montante n'a pas été bien considérée dans la littérature. De plus, pour obtenir des gains significatifs dans le MIMO-NOMA basé sur HBF, le schéma SB- et MB-NOMA doit être couplé à une méthode d'allocation optimale des ressources, y compris la sélection d'antenne, le regroupement d'utilisateurs et l'allocation de puissance, visant à optimiser les efficacités spectrale et énergétique. En outre, les canaux à ondes millimétriques étant des canaux à large bande sélectifs en fréquence, il serait intéressant d'étudier les schémas AD-DBF et AD MIMO-NOMA dans le cadre de transmissions OFDM et de proposer une allocation des ressources avec uniquement l'information angulaire.

Chapter 1

Introduction

1.1 Next-generation wireless networks

The mobile system has become an essential social infrastructure, driving the digital economy and marshaling people's everyday life. This trend requires the cellular network to link industries with the Information and Communications Technology (ICT) community and boost the user experience [1]. In addition, the Internet of Things (IoT) has emerged as a new paradigm, so massive devices and sensors will be connected to the network. In other words, the wireless system has moved from connecting people to connecting people and objects [2]. Moreover, various emerging use cases have been introduced, such as eHealth services, augmented reality/virtual reality applications, cloud computing, remote-controlled robots, automotive driving, surveillance and monitoring sensors, etc. These new use cases can be classified into three usage scenarios [3]: 1) massive machine-type communications (mMTC), 2) ultra-reliable low-latency communications (URLLC), and 3) enhanced mobile broadband (eMBB). To meet the rapid growth of data traffic, the explosion in the number of connected devices and the quality of service (QoS) requirements of mMTC, URLLC, and eMBB services, a next-generation wireless network system is called upon to develop and adopt new efficient technologies. Once the QoS requirements of the various services are met, the emerging services are capable of significantly enhancing the quality of life.

Against the above backdrop, the fifth-generation (5G) of cellular communications and beyond are rapidly emerging to address these stringent requirements, innovating promising technologies that make 5G a revolution of the previous generations rather than evolution [4,5]. In particular, the future wireless network adopts new radio spectrum as millimeter-wave (mmWave) bands, massive multiple-input multiple-output (MIMO) technology, novel multiple access techniques as non-orthogonal multiple access (NOMA), device-to-device (D2D) communications, dense cell deployment, etc. The integration of the first three technologies, namely "massive MIMO-NOMA-aided mmWave communications," has received a lot of attention in the recent literature as a new communication that can significantly improve network capacity and provide QoS for future services. The general theme of this dissertation revolves around this new type of communication.

1.1.1 MmWave communications

The spectrum in traditional cellular bands below 6 GHz is fully congested and can not satisfy the growing demands on mobile systems [6]. One promising and feasible solution is to move to the mmWave bands, which offer a vast amount of unused spectrum and achieve high data rates for future wireless network systems [1, 3]. These bands are suggested as a promising candidate for 5G and beyond cellular communications. Channel measurements reveal that mmWave signals experience significant propagation losses and are more suitable for the coverage of small cells [3].

1.1.2 Massive MIMO

Advanced multi-antenna systems, known as "massive MIMO," are considered the most promising technology to handle the exponential growth in wireless data traffic and dramatically improve spectral efficiency [7]. Specifically, massive MIMO provides a large amount of degrees-of-freedom (DoFs) and allows the cellular network to serve multiple users simultaneously with high beam-forming (BF) gains, thereby improving spectral and energy efficiencies and reliability. For instance, with massive MIMO, area throughput can be increased at least tenfold by increasing spectral efficiency (bit/s/Hz/cell) while cell density and bandwidth are the same as in real networks. These considerable gains are achieved by using hundreds of antennas at the base station (BS) to serve tens of users by spatial multiplexing [3].

1.1.3 Non-orthogonal multiple access

Next-generation wireless networks exploit novel multiple access techniques to accommodate massive connectivity and improve network capacity. These new techniques fall under the umbrella of NOMA [8, 9]. As the name implies, NOMA aims to exploit additional non-orthogonal resources and serve multiple users in the same orthogonal resource blocks (e.g., space, time, frequency). There are two different categories of NOMA, namely, power-domain NOMA (PD-NOMA) and code-domain NOMA [10–16]. Compared to orthogonal multiple access (OMA) techniques, NOMA enables massive connectivity, improves spectral efficiency and fairness among users, and reduces latency.

1.1.4 Massive MIMO-NOMA-aided mmWave communications

The integration of these technologies has been attracted a lot of attention in the literature as a novel communication significantly enhancing the network capacity [17–20]. Indeed, the corresponding tiny wavelength helps implementing massive antenna elements within a small form factor. For instance, the form factor of a 10×10 uniform rectangular array (URA) with an equal inter-element spacing distance of half-wavelength is just 5.06 cm^2 when the dedicated radio-frequency (RF) chain operates at a carrier frequency of 60 GHz. On the other hand, the massive MIMO technology provides high array gain to combat the tremendous path loss and blockage at mmWave frequencies. Moreover, NOMA improves the connectivity density by non-orthogonally scheduling the users at the same time-frequency resources. As a result, the

massive MIMO-NOMA-aided mmWave communications can manage massive connectivity and support the QoS requirements of the emerging use cases.

1.1.5 Challenges in mmWave massive MIMO-NOMA systems

Massive MIMO-NOMA systems at mmWave frequencies pose many challenging issues, including precoding design complexity, hardware constraints, power consumption, and channel estimation complexity and overhead. Indeed, the implementation of a large number of antennas is not as challenging as the cost and power consumption of digital-to-analog converters (DACs) and RF chains assigned to each antenna element. In addition, the baseband of the existing precoding techniques becomes more complex with large-scale antenna arrays (LSAAs). Moreover, most of them require a full channel-state-information (CSI), which is not easy to achieve in practice. Indeed, when massive MIMO is applied, acquiring an accurate CSI can consume excessive spectrum resources. Besides, the application of NOMA requires efficient user ordering, user clustering and power allocation techniques, which require the full CSI.

1.2 Thesis contributions

This dissertation aims to address the critical challenges of mmWave massive MIMO-NOMA systems in downlink transmissions, which result from using LSAAs, namely, the precoding design complexity, hardware constraints, and channel estimation overhead. Throughout this dissertation, we exploit the main characteristics of mmWave channels, i.e., high directionality and blockage, that make angular information a promising partial CSI, whose feedback overhead does not scale with the number of antennas. Moreover, the estimation periodicity of a transmission channel (in frequency or in time) depends intrinsically on the frequency (Doppler effect). In contrast, the estimation of the directions depends only on the mobility of each user. The main goal of this dissertation is to prove the interest of using only angular information relative to users' positions to perform low-complex precoding, user clustering and ordering, and power allocation methods for mmWave massive MIMO-NOMA systems with either digital BF (DBF) or hybrid BF (HBF) techniques. For this purpose, we applied both 2D and 3D mmWave channels using the realistic and statistical mmWave channel simulator, called NYUSIM [21–23], developed by New York University.

The main contributions of this thesis can be summarized as follows

- First, we consider mmWave massive MIMO systems with DBF techniques. To reduce the precoding design complexity and the channel acquisition overhead, we aim to analyze the spatial behavior and potentiality of the low-complex angle-domain (AD)-DBF technique in mono- and multi-path mmWave environments with LSAAs. Specifically, AD-DBF uses only digital phase shifters and requires only users' angles. As a reference, we consider the conjugate beamforming technique that maximizes the BF gain at the desired user, based on the complete knowledge of the transmission channel. In contrast to the previous work [24] that adopts a simplified mmWave channel model, we utilize the realistic and

statistical NYUSIM channel model and simulator.

- Second, we consider the AD-DBF limitation, especially in a congested cell, i.e., high inter-user interference caused by the neighboring beams, when tens of antennas are implemented at the transmitter. Instead of increasing the number of transmit antennas, thus hardware complexity and cost, we propose a NOMA-based solution that exploits the power domain and serves multiple users at the same spatial and time-frequency domains. In particular, we take advantage of the mmWave features, and we design a new AD MIMO-NOMA scheme that can handle massive connectivity in a congested scenario. In the second study, our contributions can be summarized as follows
 - We first define new angular-based metrics leveraging the main characteristics of mmWave channels, such as the user channel quality, the sum-throughput, and the inter-user spatial interference, solely based on angular information.
 - We design an AD user ordering strategy for the NOMA protocol. In contrast to the partial CSI-based strategies existing in the literature, our scheme considers inter-cluster spatial interference.
 - We develop two-user and multi-user AD clustering algorithms to reduce inter-user interference. To the best of our knowledge, our proposed algorithms are the first user clustering algorithms that use angular-based inter-user spatial interference, which requires only angular information and includes angular distance and main lobe beamwidth in a unique metric.
 - We utilize the derived angular-based sum-throughput to design a sub-optimal angle-domain power allocation technique that is consistent with the NOMA protocol in the power domain and ensures efficient SIC.
- Third, we take advantage of massive MIMO technology to implement LSAAs (e.g., with 128, 256 antennas) and increase the BF gain and thus the network capacity. However, DBF requires as many RF chains as antennas, resulting in high energy consumption and unaffordable hardware complexity and cost. For that, in this study, we shift to HBF with limited RF chains. And, we investigate AD HBF-based massive MIMO-NOMA for mmWave communications. Conventional AD MIMO-NOMA - also denoted as single-beam NOMA (SB-NOMA) - generates a single beam in the direction of the NOMA cluster. With LSAAs, the beamwidth is very narrow, and thus using the SB-NOMA scheme, only the users having similar angles can be served concurrently within the same cluster. Therefore, due to the limited RF chains in mmWave hybrid systems, SB-NOMA with LSAAs can not manage the connectivity of all users, especially in a congested cell. To solve this problem, we propose different schemes that offer additional DoFs for SB-NOMA in mmWave hybrid systems with only the knowledge of angular information. In the third study, our contributions can be summarized as follows
 - We first design a joint SB-NOMA and TDMA scheme exploiting the time-domain

- resources to ensure connectivity of all users even when the total number of clusters (single-user and SB-NOMA) is greater than that of RF chains. In contrast to the previous work [25], our proposed scheme can apply to any antenna array architecture.
- We then leverage the concept of multi-beam NOMA (MB-NOMA) to reduce the signal processing complexity when applying two different multiple access techniques. Specifically, we propose two joint SB- and MB-NOMA schemes to benefit from NOMA multi-user diversity, regardless of the cell load and user location. The former gives a significant sum-rate performance related to DBF, while the latter reduces the number of active RF chains, i.e., the power consumption.

1.3 List of publications

As a byproduct of the above contributions, this dissertation has led to the following publications

1.3.1 Journal publication

- **I. KHALED**, C. LANGLAIS, A. EL FALOU, B. ELHASSAN, M. JEZEQUEL, "Multi-user angle-domain MIMO-NOMA system for mmWave communications", IEEE Access, 2021, vol. 9, p. 129443-129459.

1.3.2 Conference publications

- **I. KHALED**, C. LANGLAIS, A. EL FALOU, M. JEZEQUEL, B. ELHASSAN, "Joint Single-and Multi-beam angle-domain NOMA for hybrid mmWave MIMO systems", WSA 2021; 25th International ITG Workshop on Smart Antennas. VDE, 2021.
- **I. KHALED**, C. LANGLAIS, A. EL FALOU, M. JEZEQUEL, B. ELHASSAN, "Joint SDMA and power-domain NOMA system for multi-user mm-Wave communications", 2020 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2020.
- **I. KHALED**, A. EL FALOU, C. LANGLAIS, B. ELHASSAN, M. JEZEQUEL, "Multi-user digital beamforming based on path angle information for mm-wave MIMO systems" WSA 2020; 24th International ITG Workshop on Smart Antennas. VDE, 2020.
- **I. KHALED**, A. EL FALOU, C. LANGLAIS, B. ELHASSAN, M. JEZEQUEL, "Performance evaluation of linear precoding mmwave multi-user MIMO systems with NYUSIM channel simulator", 2019 2nd IEEE Middle East and North Africa COMMUNICATIONS Conference (MENACOMM). IEEE, 2019.

Next, we give an overview of the organization of this manuscript.

1.4 Organization of the thesis

The rest of this dissertation is organized as follows

- **Chapter 2** provides a general overview of the topics relevant to this dissertation. First, we present the basic theory of the uniformly excited antenna array, massive MIMO, and NOMA. Then, we review the various channel models and simulators presented in the literature, especially mmWave channel models and simulators. Next, we detail the mmWave statistical spatial channel model implemented in NYUSIM, which will be used throughout this dissertation.
- **Chapter 3** analyzes the spatial behavior and potentiality of AD-DBF for mmWave massive MIMO systems. This analysis has been performed using NYUSIM and considering as a reference the conjugate beamforming, the DBF technique well known in the literature.
- **Chapter 4** reviews existing MIMO-NOMA schemes with full or partial CSI. Then, we introduce a new AD MIMO-NOMA scheme for mmWave communications. Specifically, we leverage the main features of mmWave channels and propose new angular-based metrics to design and develop effective user ordering, user clustering, and power allocation techniques. In this chapter, the generalization ability is investigated. We study not only the case of two-user NOMA groups, but also the case of multi-user NOMA groups.
- **Chapter 5** reviews existing MIMO-NOMA designed based on HBF techniques. Then, we study the AD-HBF-based MIMO-NOMA systems, where the BS deploys LSAAs. And, we propose different schemes to handle multi-user connectivity even in an overloaded scenario, i.e., the number of users is greater than that of RF chains. These solutions benefit from TDMA or MB-NOMA to provide additional DoFs for SB-NOMA with LSAAs in mmWave hybrid systems.
- **Chapter 6** summarizes the work presented in this dissertation and discusses some interesting directions for future research in the scope of this dissertation.

Chapter 2

Background

2.1 Introduction

In this chapter, we introduce the main topics relevant to the work presented throughout this dissertation. First, we give the main features of equally spaced linear and planar arrays with uniform excitation. Then, we provide the principal aspects of conventional and massive MIMO systems for downlink transmissions. And, we classify the different types of multi-user MIMO beamforming techniques. Next, we give the fundamentals of the downlink power-domain NOMA transmission. Finally, a thorough review of the sub-6 GHz and mmWave channel models is provided. And, the realistic and statistical mmWave channel model simulator, called NYUSIM, is detailed and will be considered throughout this dissertation.

2.2 Uniformly excited phased arrays

The antenna, which transmits and receives signals, is one of the most essential components for wireless communications [26,27]. And, the implementation of antenna arrays has attracted much attention in wireless communications to maximize the radiation power in a specific direction [28–30]. For instance, a phased antenna array consists of multiple antenna elements; each one is fed coherently and uses a time-delay or a variable excitation control to generate a beam in the scan angle. The antenna excitation control includes both amplitude and phase controls. Variable amplitude control, which affects the pattern shaping, is not always provided. For the uniformly excited phased array, all antennas have a uniform excitation [28–30]. Throughout this dissertation, we assume an equal spacing distance and, we consider the excitations that are constant in amplitude but with uniform scan phases. To this end, we discuss, in this section, the basic characteristics of equally and uniformly excited phased arrays.

Let's consider a planar array having $M = M_u \times M_v$ antennas and located in the uov plane, with $u \neq v \in \{x, y, z\}$. Denote M_u and M_v as the number of antennas along the u -axis and the v -axis, respectively.

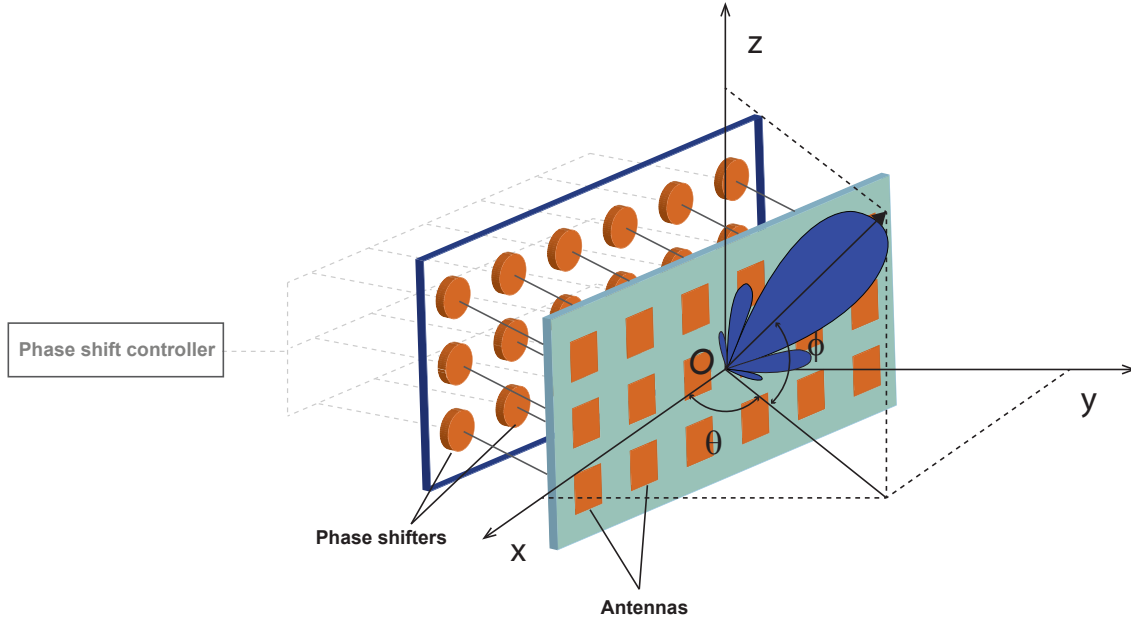


Figure 2.1: Representation of the coordinate system and the phased planar array located in the xoz plane.

2.2.1 Array steering vector

Let $\mathbf{v}$ be the vector representing the phase delay experienced by a plane wave at each antenna relative to a given reference, which is assumed to be the first antenna element. More precisely, $[\mathbf{v}(\vec{\Theta})]_{m_u, m_v}$ is given by $[\mathbf{v}(\vec{\Theta})]_{m_u, m_v} = e^{j\iota_{m,n}}$, where $\iota_{m,n}$ is the spatial phase delay experienced by the signal arriving or departing from the $\vec{\Theta}$ angle at the (m_u, m_v) -th antenna w.r.t. the first element. Fig. 2.1 illustrates the coordinate system used throughout this thesis, where θ and ϕ are the azimuth and the elevation angles, respectively. Denote $\vec{\Theta} = (\theta, \phi)$ as the couple of these angles. Therefore, $[\mathbf{v}(\vec{\Theta})]_{m_u, m_v}$ can be expressed as follows

$$[\mathbf{v}(\vec{\Theta})]_{m_u, m_v} = e^{j\{(m_u-1)\omega_u(\vec{\Theta}) + (m_v-1)\omega_v(\vec{\Theta})\}} \quad (2.1)$$

where $j \triangleq \sqrt{-1}$ is the imaginary unit, and $\omega_u(\vec{\Theta})$ is the inter-element spatial phase delay along the u -axis and is given by

$$\omega_u(\vec{\Theta}) = \begin{cases} 2\pi \frac{d}{\lambda} \cos(\theta) \cos(\phi) & \text{if } u = x, \\ 2\pi \frac{d}{\lambda} \sin(\theta) \cos(\phi) & \text{if } u = y, \\ 2\pi \frac{d}{\lambda} \sin(\phi) & \text{if } u = z. \end{cases} \quad (2.2)$$

Denote λ as the carrier wavelength and d as the inter-element spacing distance. Throughout the rest of this thesis, we assume an inter-element spacing distance of half-wavelength, i.e., $d = \frac{\lambda}{2}$. Fig. 2.2 shows the inter-element path difference of the signal propagation along the axis in which the antennas are stacked.

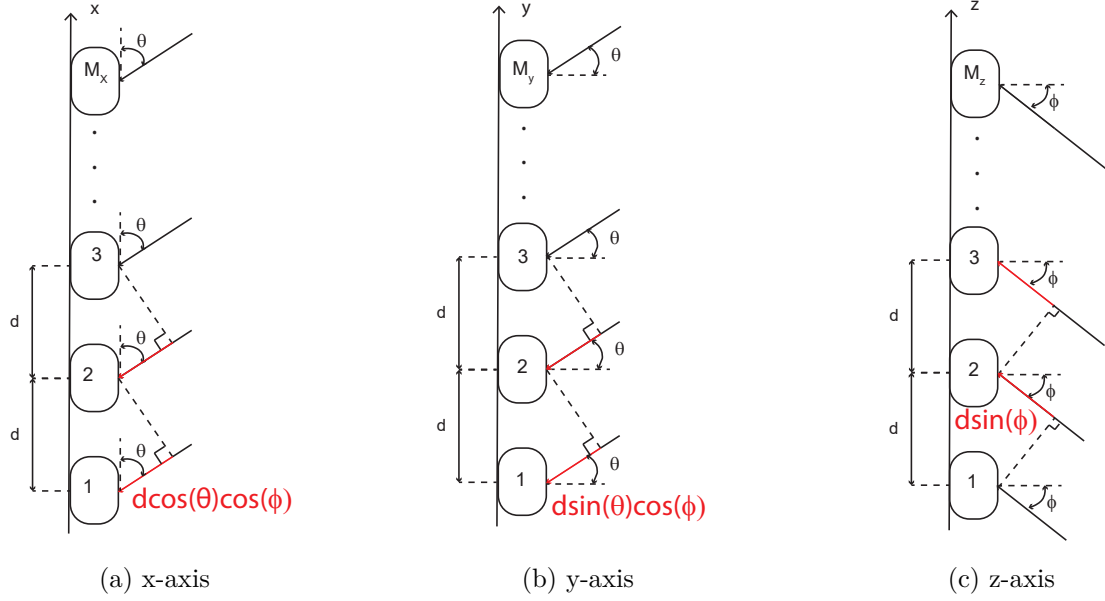


Figure 2.2: Representation of the inter-element path difference along x-, y-, and z-axis experienced by the plane wave upcoming from the $\vec{\Theta}$ angle w.r.t the antenna array.

To steer a beam in the $\vec{\Theta}$ scan angle, the complex weight excitation $[\mathbf{a}(\vec{\Theta})]_{m_u, m_v}$ assigned to the (m_u, m_v) -th antenna is designed to cancel the phase shifts at the related antenna, thereby is given by $[\mathbf{a}(\vec{\Theta})]_{m_u, m_v} = [\mathbf{v}(\vec{\Theta})]_{m_u, m_v}^*$. Hence, the array steering vector $\mathbf{a}(\vec{\Theta})$ has the effect of moving each antenna's output to the $\vec{\Theta}$ scan angle, therefore focusing the beam.

By referring to (2.1), the steering vector $\mathbf{a}(\vec{\Theta}) \in \mathbb{C}^{M \times 1}$ of uniform planar array can be expressed as follows

$$\mathbf{a}(\vec{\Theta}) = \left[1, e^{-j\omega_u(\vec{\Theta})}, \dots, e^{-j(M_u-1)\omega_u(\vec{\Theta})}, e^{-j\{\omega_u(\vec{\Theta})+\omega_v(\vec{\Theta})\}}, \dots, e^{-j\{(M_u-1)\omega_u(\vec{\Theta})+(M_v-1)\omega_v(\vec{\Theta})\}} \right]^T = \mathbf{a}_u(\vec{\Theta}) \otimes \mathbf{a}_v(\vec{\Theta}), \quad (2.3)$$

where $\mathbf{a}_u(\vec{\Theta})$ and $\mathbf{a}_v(\vec{\Theta})$ are the steering vectors of the M_u - and M_v -elements uniform linear array (ULA) stacked along the u - and v -axis, respectively, and are given by

$$\mathbf{a}_u(\vec{\Theta}) = \left[1, e^{-j\omega_u(\vec{\Theta})}, \dots, e^{-j(M_u-1)\omega_u(\vec{\Theta})} \right]^T, \quad (2.4)$$

$$\mathbf{a}_v(\vec{\Theta}) = \left[1, e^{-j\omega_v(\vec{\Theta})}, \dots, e^{-j(M_v-1)\omega_v(\vec{\Theta})} \right]^T. \quad (2.5)$$

2.2.2 Array factor and radiation pattern

Array factor

The *array factor* AF is a key function to characterize an antenna array. AF depends on different parameters, such as the array architecture (e.g., linear, circular, rectangular, spherical, etc.), the array size, the inter-element spacing distance, the excitation amplitude, and the excitation phase of the individual elements.

When an equally spaced and uniformly excited planar array scans a beam to the $\vec{\Theta}_0 = (\theta_0, \phi_0)$

angle, the set of excitation at the different antennas is given by $[\mathbf{a}(\vec{\Theta}_0)]$ to compensate the spatial phase delays and maximize the beamforming gain at the steering direction $\vec{\Theta}_0$. Hence, the corresponding array factor $\text{AF}_{\vec{\Theta}_0}(\vec{\Theta})$ can be expressed as follows [28, 29]

$$\text{AF}_{\vec{\Theta}_0}(\vec{\Theta}) = \sum_{m_v=1}^{M_v} \sum_{m_u=1}^{M_u} [\mathbf{a}(\vec{\Theta}_0)]_{m_u, m_v} e^{j\{(m_u-1)\omega_u(\vec{\Theta}) + (m_v-1)\omega_v(\vec{\Theta})\}}, \quad (2.6)$$

where $[\mathbf{a}(\vec{\Theta}_0)]_{m_u, m_v}$ is the complex weight assigned to the (m_u, m_v) -th antenna element to point a beam in the $\vec{\Theta}_0$ angle. Thereby, (2.6) can be rewritten by

$$\text{AF}_{\vec{\Theta}_0}(\vec{\Theta}) = \sum_{m_v=1}^{M_v} \sum_{m_u=1}^{M_u} e^{j\{(m_u-1)(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0)) + (m_v-1)(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\}}, \quad (2.7a)$$

$$= \sum_{m_v=1}^{M_v} e^{j\{(m_v-1)(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\}} \sum_{m_u=1}^{M_u} e^{j\{(m_u-1)(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))\}}. \quad (2.7b)$$

$$\stackrel{(a)}{=} \frac{1 - e^{jM_v(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))}}{1 - e^{j(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))}} \frac{1 - e^{jM_u(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))}}{1 - e^{j(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))}}, \quad (2.7c)$$

where (a) is obtained using the sum of a geometric series, i.e., $\sum_{n=1}^N q^{N-1} = \frac{1-q^N}{1-q}$.

If we consider only the magnitude of the array factor, this becomes [28, 29]

$$\text{AF}_{\vec{\Theta}_0}(\vec{\Theta}) = \frac{\sin\left(\frac{M_v}{2}(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\right)}{\sin\left(\frac{1}{2}(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\right)} \frac{\sin\left(\frac{M_u}{2}(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))\right)}{\sin\left(\frac{1}{2}(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))\right)}, \quad (2.8)$$

Normalized to its peak value given by $\text{AF}_{\vec{\Theta}_0}(\vec{\Theta}_0) = M_u \times M_v = M$, this implies

$$\text{AF}_{\vec{\Theta}_0}^n(\vec{\Theta}) = \frac{\sin\left(\frac{M_v}{2}(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\right)}{M_v \sin\left(\frac{1}{2}(\omega_v(\vec{\Theta}) - \omega_v(\vec{\Theta}_0))\right)} \frac{\sin\left(\frac{M_u}{2}(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))\right)}{M_u \sin\left(\frac{1}{2}(\omega_u(\vec{\Theta}) - \omega_u(\vec{\Theta}_0))\right)}. \quad (2.9)$$

Array radiation pattern

By definition, the *antenna radiation pattern* or *antenna pattern* is the graphical representation of the antenna radiation characteristics as a function of space coordinates [26]. In other words, the antenna's pattern represents how the antenna radiates energy out into space. The *array radiation pattern* or *array pattern* AP_{array} is the product of the array factor AF and the individual antenna pattern AP_{element} as follows [26, 28, 29]

$$AP_{\text{array}} = AP_{\text{element}} \times AF. \quad (2.10)$$

Fig. 2.3 illustrates the relation in (2.10), where the left figure represents the individual antenna pattern AP_{element} , the center figure represents the array factor AF and the right figure plots the corresponding array pattern, i.e., the product of both of them. Note that the array is a ULA array constructed using four directive antennas [24].

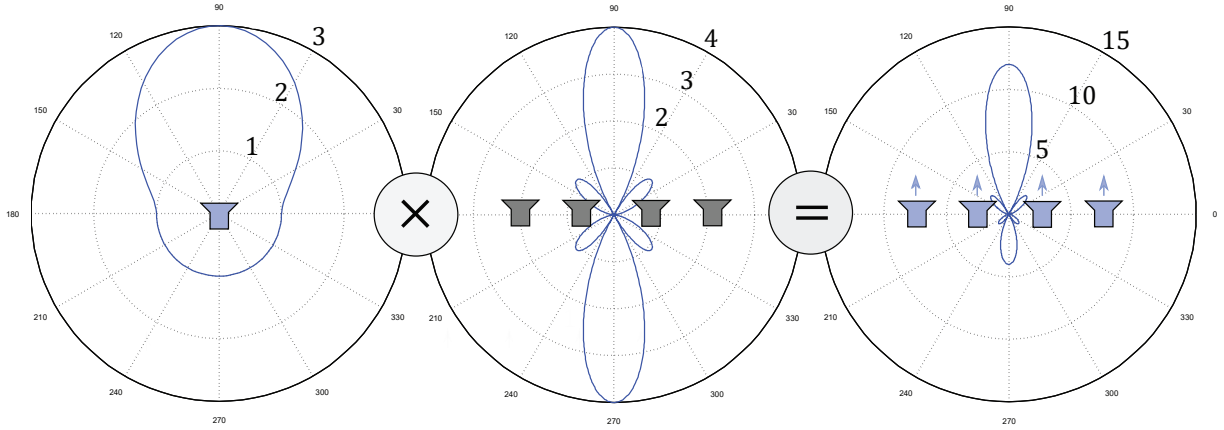


Figure 2.3: Illustration of the radiation pattern resulting from the product of the individual antenna pattern and the array factor of an array with four directive antennas. The figure is taken from [24].

The array pattern consists of different lobes that are sub-classified into main, side, minor, and back lobes [26]. Specifically, a radiation lobe is part of the array pattern at which the radiated signal intensity reaches a maximum, separated by regions of relatively weak radiation intensity or nulls. Some have more significant radiation intensity than others, but all are sub-classified as lobes. The main lobe - also known as the main beam - is the radiation lobe containing the direction of the greater radiation intensity. In Fig. 2.3, the main lobe is scanned in the $\theta = 90^\circ$ direction. Minor lobes represent the radiation in undesired directions, i.e., any other lobe from the main lobe. The minor lobe level is the ratio of the maximum radiation intensity in the lobe in question to that in the main lobe. Side lobes are normally the minor lobe having the greatest radiation intensity. A back lobe is the opposite radiation lobe w.r.t the main beam.

With *isotropic antenna* - also known as *omnidirectional antenna*, the signal is uniformly transmitted in all directions [31]. In other words, the received power signal P_R is the same at any direction for a given distance d_R from the antenna element, as shown in Fig. 2.4 and, in a free space, is given by

$$P_R = \frac{P_0 \lambda^2}{(4\pi d_R)^2}, \quad (2.11)$$

where P_0 is the transmitted power.

Such antenna has a gain of 1 (0dB) in the spherical space all around it, and thereby its corresponding normalized pattern is equal to 1, i.e., $AP_{\text{element}}^n = 1$. Accordingly, using isotropic antennas, the normalized array pattern AP_{array}^n will be equal to the normalized array factor AF^n , i.e., $AP_{\text{array}}^n = AF^n$. Throughout this dissertation, we only consider isotropic antenna elements, thus AP_{array}^n and AF^n are equivalent.

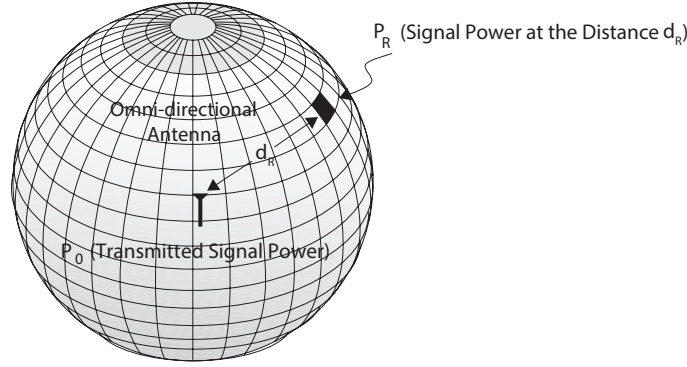


Figure 2.4: Omnidirectional antenna propagation in a free space [31]

2.2.3 Beamwidth

For a uniform phased array, the half-power points are the points at which the radiation intensity has dropped to half (-3dB) of its peak value. In other words, these points are found by setting $AP_{(\theta_0, \phi_0)}^n(\theta, \phi) = \sqrt{0.5}$, where $AP_{(\theta_0, \phi_0)}^n(\theta, \phi)$ is the normalized array pattern corresponding to a uniform phased $M_u \times M_v$ array that scans a beam in the $\vec{\Theta}_0$ direction. The *half-power beamwidth* $\vec{\Omega}^{3\text{dB}} = (\Omega_{\text{az}}^{3\text{dB}}, \Omega_{\text{el}}^{3\text{dB}})$ - also called as *3dB-beamwidth* - is the angular distance between the half-power points of the main lobe, as shown in Fig. 2.5. Denote $\Omega_{\text{az}}^{3\text{dB}}$ and $\Omega_{\text{el}}^{3\text{dB}}$ as the azimuth and the elevation widths of the beam scanned in the $\vec{\Theta}_0$ direction, respectively.

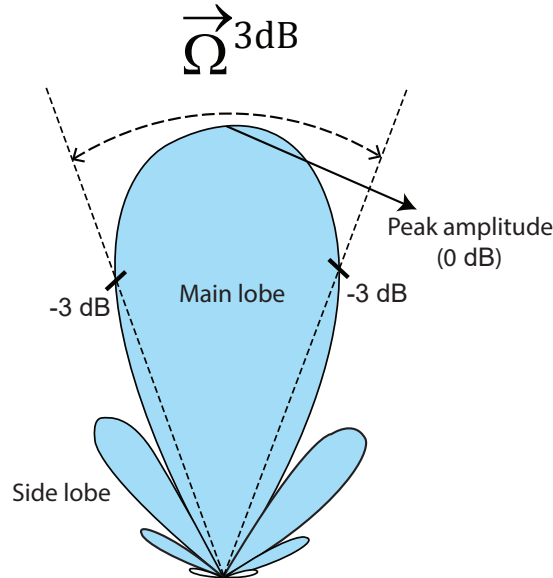


Figure 2.5: Illustration of 3dB-beamwidth

Fig. 2.6 plots the normalized antenna pattern for a ULA along the x -axis, i.e., $M_u = M_x$, $M_v = 1$, $M = M_x$ and $\vec{\Theta} = (\theta, 0)$. In Fig. 2.6, we vary the number of transmit antennas $M \in \{16, 32, 64\}$ and the scan angle $\theta_0 \in \{30^\circ, 60^\circ\}$. It can be seen that the 3dB-beamwidth depends on two factors, the number of transmit antennas M and the scan angle θ_0 . The more

antennas, the narrower the beam.

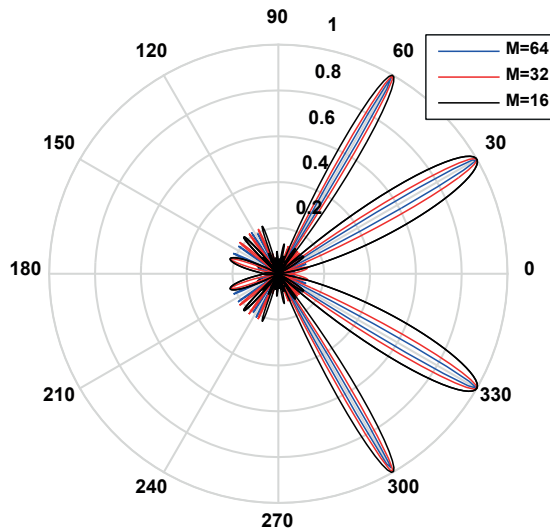


Figure 2.6: Normalized pattern of a ULA array having $M \in \{16, 32, 64\}$ antenna elements and generating a beam at the azimuth angle $\theta_0 \in \{30^\circ, 60^\circ\}$.

Fig. 2.7 plots the normalized antenna pattern for uniform rectangular array (URA) located in the xoz plane, i.e. $u = x$ and $v = z$. In Fig. 2.7, we fix the number of horizontal antennas $M_x = 32$ and the scan angle $\vec{\Theta}_0 = (60^\circ, 30^\circ)$, and we vary the number of elevation antennas $M_z \in \{4, 6\}$. It can be seen that the elevation beamwidth becomes narrower when more elevation antennas are implemented.

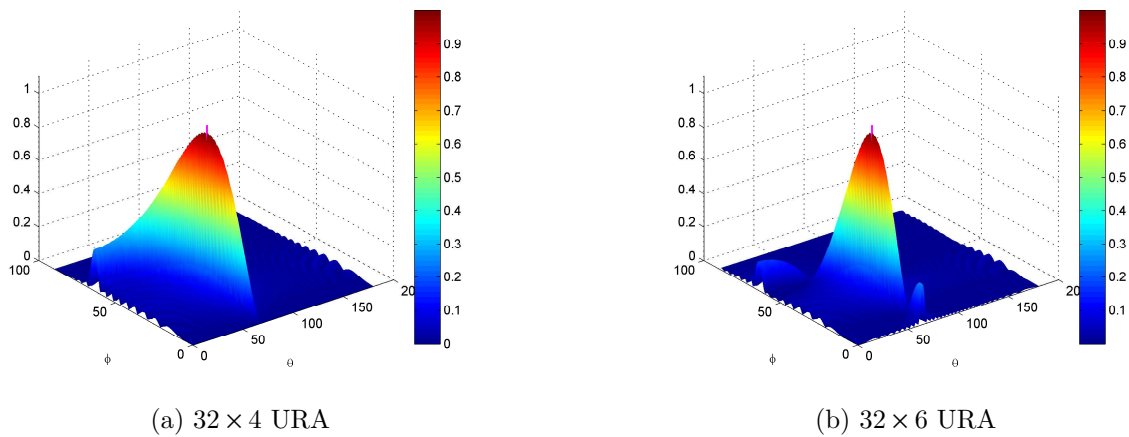


Figure 2.7: Normalized 3D array pattern of a beam scanned at $\vec{\Theta}_0 = (60^\circ, 30^\circ)$ with (a) 32×4 URA and (c) 32×6 URA.

2.3 Key aspects of massive MIMO systems

Throughout this section, we explain the interest of enhancing the spectral efficiency in a wireless network and how MIMO, especially m-MIMO, can achieve this goal by reviewing the different benefits of using multiple antennas. And, we present the different types of beamforming tech-

niques based on signal processing and hardware implementations. Finally, we investigate the challenges arising from the application of LSAAs.

2.3.1 Enhancing the spectral efficiency

The area throughput of a wireless network [$bit/s/km^2$] can be expressed as follows [32, 33]

$$\text{Area throughput}[bit/s/km^2] = \text{Spectral efficiency}[bit/s/Hz] \times \text{Bandwidth}[Hz] \times \text{Cell density}[cells/km^2] \quad (2.12)$$

From (2.12), three key components can yield higher area throughput: i) improved data transmission efficiency (for a given bandwidth and per cell), ii) more allocated spectrum, and iii) network densification by increasing the number of access points.

The previous cellular generations aimed to improve the area throughput by focusing on two factors; allocating more bandwidth and applying cell densification [33]. Nowadays, especially in urban environments where data traffic demands are the highest, wireless local area networks are available almost everywhere, and the inter-site distance in the cellular network is about a few hundred meters. Cell densification is still possible, but it looks like it attends a saturation point. In addition, the frequency spectrum below 6 GHz are highly congested. Recently, for 5G and beyond, the spectrum above 6 GHz is attracting much attention because it offer large amounts of bandwidth. However, these bands may only perform well under short-range LoS conditions due to significant path loss. Instead, the SE factor has not seen substantial improvements in previous cellular generations. Hence, the 5G and beyond networks aim to achieve high area throughput by significantly improving the SE factor. One possible solution is the suggestion of massive MIMO, which can enhance the SE by one or two orders of magnitude, as a fundamental technology for future wireless networks [7, 32, 33]. Next, we give the different benefits of using multiple antennas in wireless communications.

2.3.2 Benefits of multiple antennas in wireless communications

In a single-input single-output (SISO) communication, a single-antenna transmitter (Tx) communicates with a single-antenna receiver (Rx). Through the additive white Gaussian noise (AWGN) channels, the SE of such systems is upper bounded by the Shannon capacity [33]. The latter is equal to $\log_2(1 + \text{SNR})$, where SNR denotes the signal-to-noise ratio. It is clear that an exponential increase in SNR, i.e., an exponential increase in the power of the transmitted signal, will lead to a linear increase in SE. Therefore, the cellular networks need to look for another way to improve the SE rather than exponentially increasing the transmitted power.

In cellular networks, the BS serves multiple users. Traditionally, the BS splits the orthogonal radio resources (e.g., time, frequency) into several resource blocks, with each block accommodating only one user. In SISO mode, the BS can transmit a single stream, and each user receives a stream with a $\text{SE} = \log_2(1 + \text{SNR})$ at each resource block. The parallel transmission, i.e., transmitting multiple streams in the same block, is considered an efficient way to enhance the SE. For example, when the BS transmits independently N_s data streams at the same block,

the sum SE can be expressed as $N_s \log_2(1 + \text{SNR})$. This improvement in SE has been achieved by adopting the multiple antennas technology - also known as multiple-input multiple-output (MIMO) [34, 35].

There are two distinct MIMO categories:

- Point-to-point MIMO - also known as single-user MIMO (SU-MIMO) - emerged in the late 1990s and is the simplest form of MIMO [36–40]: Within the same block of time-frequency resources, a single user with multiple antennas communicates with the BS with multiple antennas, as shown in Fig. 2.8a, where each user, i.e., time-frequency resource, is represented by a color. The different users are orthogonally multiplexed, e.g., in time and frequency domains using time-division multiple access (TDMA) and frequency-division multiple access (FDMA), respectively. From Fig. 2.8a, we see that the BS radiates energy uniformly throughout the cell, which is not very energy efficient.
- Multi-user MIMO (MU-MIMO) [41–44]: The BS equipped with an antenna array serves multiple users in the same time-frequency resources. Specifically, the BS performs space division multiple access (SDMA) by applying a beamforming technique to separate users in the spatial domain; see Fig. 2.8b. Interestingly, such systems allocate the entire bandwidth and overall symbol time to each user, and focus the energy in the direction of the user. It should be noted that MU-MIMO is less sensitive to the propagation channel and works well with only single-antenna terminals, unlike to SU-MIMO [7]. Therefore, without loss of generality, our work focuses on MU-MIMO and considers single-antenna user terminals to reduce the end-user products and the power consumption of battery.

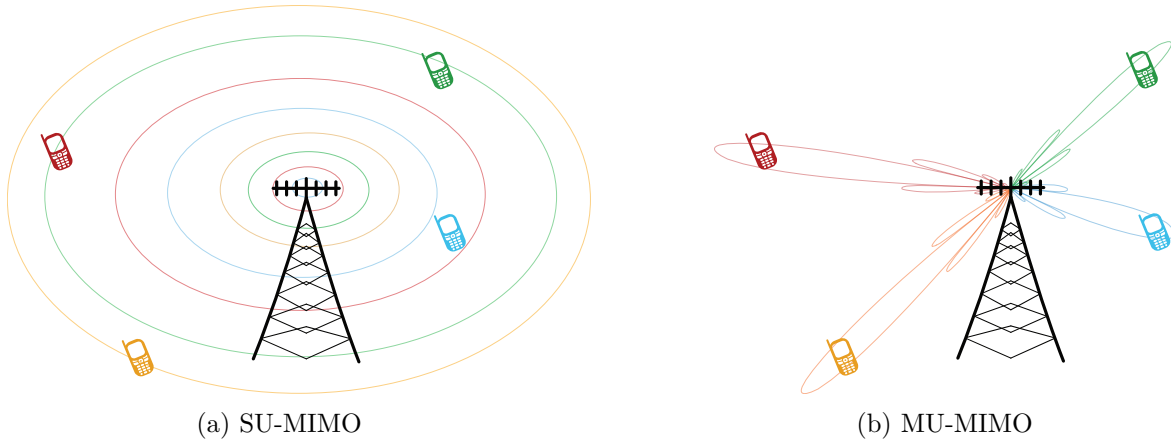


Figure 2.8: Illustration of (a) SU-MIMO and (b) MU-MIMO systems in downlink [24].

The use of multiple antennas in wireless communications can improve both the SE and reliability of link transmission by yielding different benefits, such as beamforming, spatial multiplexing, and spatial diversity gains.

1. Beamforming gain

Using a single isotropic antenna, the BS transmits the desired signal in all directions.

Hence, the power is emitted everywhere. It is interesting to note that the use of multiple transmit antennas allows the base station to form a directional beam. More antennas lie at the BS, more the main lobe is directive, and more the transmit power is focused in one direction. Therefore the SNR increases and so does the SE. The *array* or *beamforming gain* is the gain in received signal power achieved by using multiple antennas, w.r.t. a single antenna. Note that the attainable array gain is proportional to the number of antennas.

2. Spatial multiplexing gain

MIMO exploits new degrees of freedom (DoFs) in the spatial domain and allows the transmission of multiple data streams assigned to one or more users at the same time-frequency resources. These data streams are spatially separated by the use of multiple antennas at Tx. This MIMO feature is called *spatial multiplexing gain*. Note that the SE increases when more data streams are served in the same time-frequency resources, and the achievable multiplexing gain is proportional to the number of data streams.

3. Spatial diversity gain

In uplink (UL) transmission and in a severe multipath fading environment, the BS exploits the *spatial* or *antenna diversity*, resulting from the presence of multiple antennas to combat the fading channel and extracts a more robust low fading signal. In effect, the UL antennas at the BS observe different fading realizations. While beamforming and multiplexing gains increase the SE, the diversity gain can improve the reliability of the data reception.

2.3.3 Multi-user MIMO beamforming

To further improve the SE of a MU-MIMO system, i.e., the signal-to-interference plus noise ratio (SINR), the BS adopts the precoding technique - also known as beamforming - before the data transmission. Based on the prior knowledge of the channel state information (CSI), the precoding technique multiplexes the different data streams at each antenna element in order to improve the received power and reduce the inter-user interference, thus increasing the SE; see Fig. 2.9. Interestingly, the rich DoFs, due to the presence of more antennas at the BS, provide a great precoding flexibility to separate users in space.

Depending on the hardware implementation and signal processing at the base station, the beamforming technique can be classified into the three following categories

- RF analog beamforming,
- baseband digital beamforming,
- hybrid RF-baseband beamforming.

Fig. 2.10 illustrates the block diagram of each beamforming technique.

1. RF analog beamforming

Analog beamforming (ABF) connects a single RF channel to all antennas via RF precoding; see Fig. 2.10a. Thus, only one signal can be transmitted. First, baseband modulation

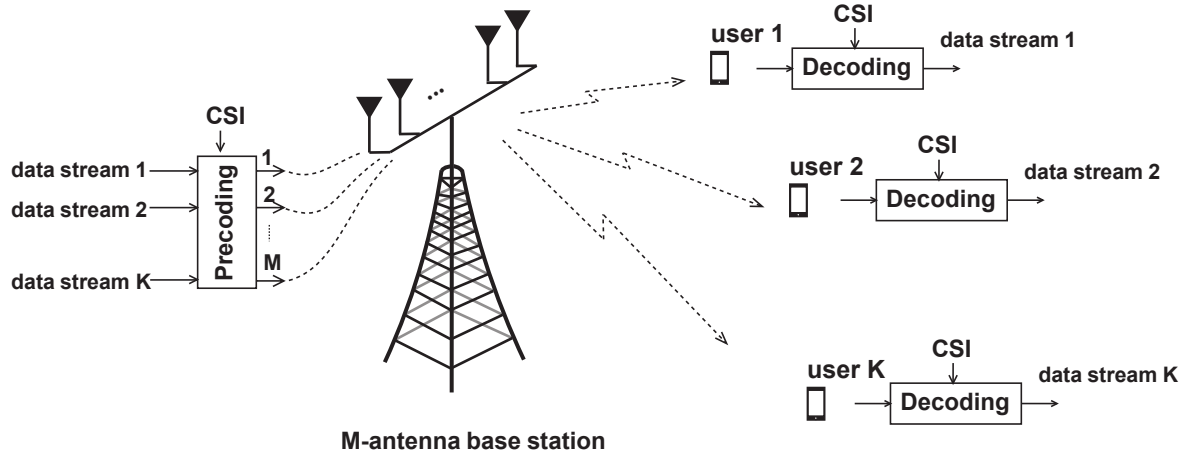


Figure 2.9: Downlink MU-MIMO system.

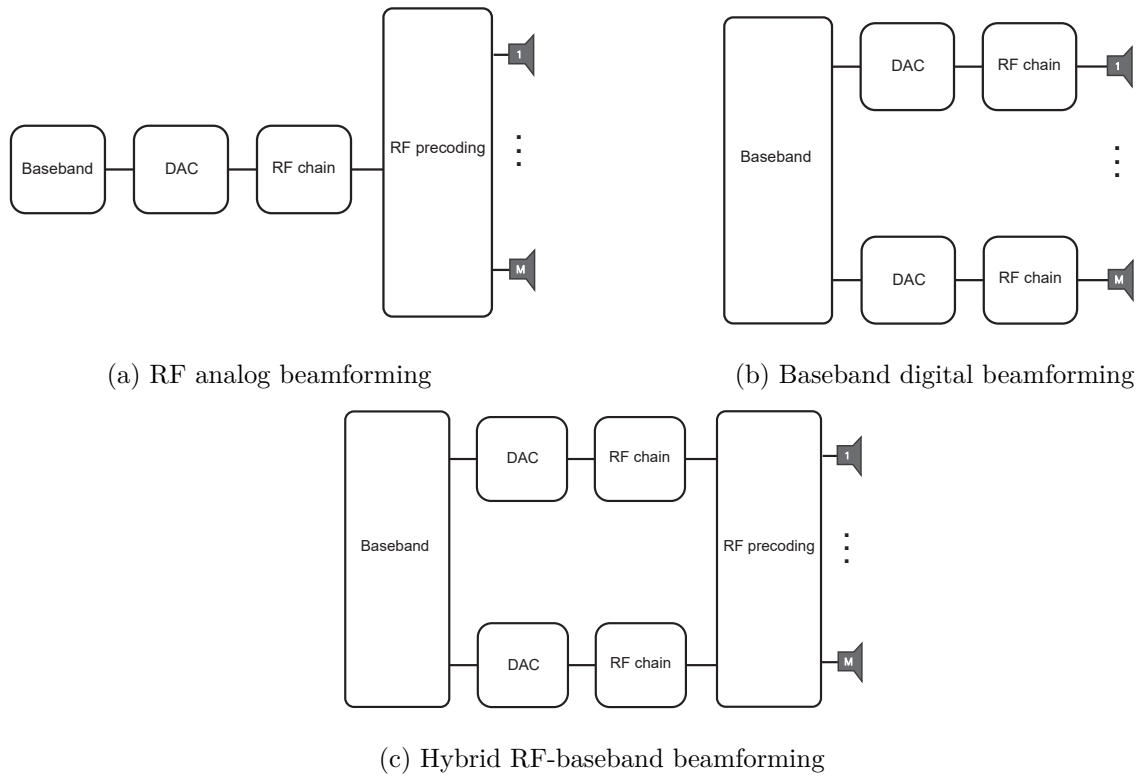


Figure 2.10: Block diagram of the (a) RF analog beamforming, (b) baseband digital beamforming and (c) hybrid RF-baseband beamforming technique.

modulates the baseband signal to the desired RF carrier frequency. Then, the RF precoding technique adjusts the amplitude and phase of the continuous RF signal, using a power amplifier and an analog phase shifter. Because only one RF chain is available at Tx, this type of beamforming allows for low hardware complexity and low power consumption, even with LSAs. However, the application of ABF is limited to low-rate transmissions because a single RF chain cannot exploit spatial multiplexing [45]. In addition, finite analog shifters restrict the flexibility of beamforming directions.

2. Baseband digital beamforming

An alternative solution to the limitations of ABF is the baseband digital beamforming (DBF), which adjusts the weight of each antenna element at the baseband. The DBF technique is more flexible to the channel variations due to the gain provided by digital signal processing. In addition, it offers high DoFs to exploit the spatial multiplexing since each antenna is connected to a dedicated RF chain; see Fig. 2.10b. However, these advantages over ABF are achieved at the expense of high energy consumption and high hardware complexity and cost, especially with LSAAAs.

3. Hybrid RF-baseband beamforming

ABF can generate a high BF gain with LSAAAs, but cannot exploit spatial multiplexing. On the other hand, DBF offers high DoFs and provides better performance and flexibility than ABF, at the expense of high energy consumption, hardware complexity, and cost. Alternatively, hybrid RF-baseband beamforming (HBF) is a beamforming technique that strikes a balance between performance, power consumption, hardware complexity, and cost. HBF, which includes both fully- and sub-connected structures, separates the signal processing into a low-dimensional baseband DBF (performed by a small number of RF chains) and a high-dimensional ABF in the RF band to increase the array gain; see Fig. 2.10c. Specifically, for hybrid fully-connected structures, each RF chain is connected to all antennas; see Fig. 2.11a. For hybrid sub-connected structures, each RF chain is connected to only a disjoint subset of antennas; see Fig. 2.11b.

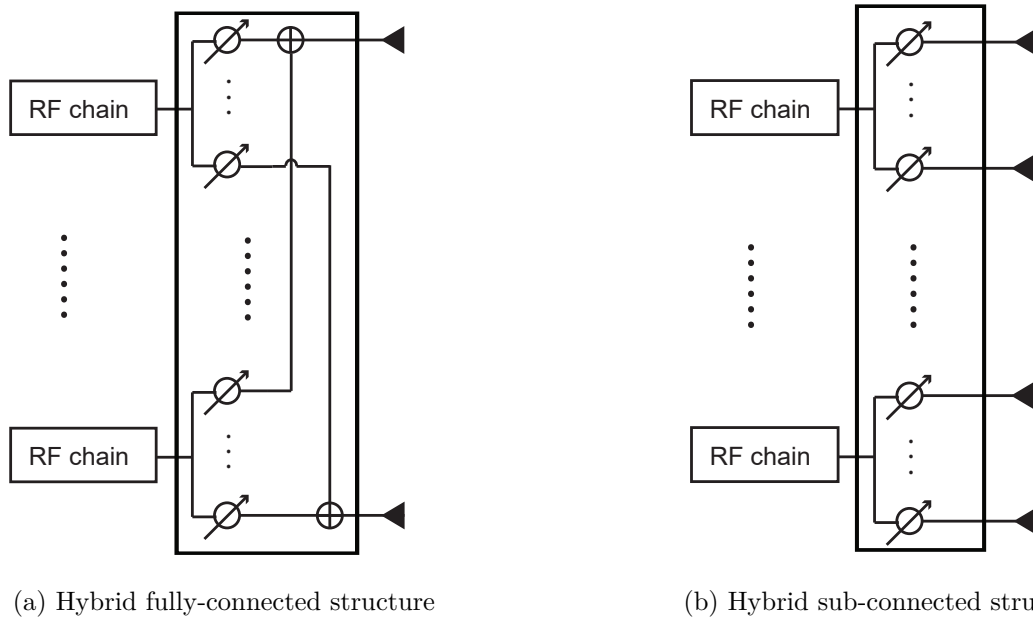


Figure 2.11: RF precoding for hybrid (a) fully-connected and (b) sub-connected structures.

2.3.4 From conventional to massive MIMO

As mentioned earlier, instead of increasing the transmitted power, the MIMO technology can exploit the beamforming gain, the spatial multiplexity, and diversity gains to improve the reliability of data reception and increase the SE, thus improving the area throughput. Moreover, the SE scales up with the number of antennas. MU-MIMO or SDMA was introduced in 4G, where the BS serves multiple users simultaneously with the same order as the number of transmit antennas. To meet the ever-increasing of data traffic and the future demands, Marzetta [7, 46–48] has shown the efficiency of implementing a large number of transmit antennas, $M \gg 1$, to serve a finite number of users, K , at the same time-frequency resources. This design is commonly referred to as m-MIMO or LSAAs. The resulting DoF richness provides great flexibility in beamforming to achieve spatial separation at the Tx and Rx. Using appropriate precoding techniques, the effects of interference and noise are negligible as the number of antennas approaches infinity, and thereby m-MIMO can produce high gains in energy efficiency (EE) and SE compared to conventional MIMO systems [49]. In addition, low-complexity linear precoders with m-MIMO have a high-performance gain thanks to the ability of a large number of antennas to suppress interference and concentrate radiated energy [50, 51].

The effectiveness of m-MIMO depends mainly on the availability of CSI at Tx (CSIT) [32, 49]. This is because the BS needs to know the DL channel of each user to perform multi-user beamforming. However, a direct estimation of the user's DL channel is not practical when the BS adopts a massive number of antennas due to the large number of pilots to be transmitted and the amount of channel information to be sent back to the BS. Interestingly, in the time division duplexing (TDD) mode, the UL and DL channels can be assumed to be identical since the UL and DL transmissions operate in the same frequency band with separate times. For this purpose, the BS can exploit the reciprocity of the propagation channel, and derives the DL channel from the estimated UL channel. Furthermore, the estimation time required by TDD to estimate the CSI depends only on the number of users, K , whereas the time required by frequency division duplexing (FDD) depends on the excessive number of transmit antennas, M [32, 49]. For this reason, TDD is a preferred operation mode over FDD for m-MIMO systems.

Conventionally, m-MIMO systems shall operate in TDD mode using linear UL and DL processing [7].

2.4 Non-orthogonal multiple access

In this section, we review the different multiple access techniques, and we describe the fundamentals of power-domain NOMA in downlink transmissions.

Multiple access (MA) is one of the critical topics for cellular network systems. The MA technique enables users to share the available resources on a single channel. Cellular networks from 1G to 4G have applied one or more of the following MA techniques: TDMA, FDMA, Code Division MA (CDMA), SDMA, and Orthogonal Frequency Division MA (OFDMA). These techniques refer as orthogonal MA (OMA) because they allow for orthogonal multiplexing of users.

In other words, orthogonal radio resources (e.g., time, frequency, space or code) are divided into sub-resources that ensure minimal inter-group interference and shared among users. Recently, as mentioned in Chapter 1, NOMA has been suggested as a potential solution to provide massive connectivity over limited resources [9]. Specifically, NOMA enables the transmission of different data streams over the same radio resources and uses signal detection methods at the Rx side to decode the multiplexed data streams. In general, many NOMA techniques are considered in the literature and fall into two categories: i) power-based NOMA [10, 11, 52] (also known as PD-NOMA) and ii) code-based NOMA. The latter introduces redundancy via coding/spreading to facilitate the users separation at Rx. It includes sparse code MA (SCMA) [12], low-density spreading (LDS) [13–15], pattern division MA (PDMA) [16], etc. This thesis focuses on the PD-NOMA technique that has been mainly adopted for downlink transmissions, with more details given later in this section. Note that for simplicity, we refer to the PD-NOMA technique as NOMA throughout this dissertation.

The principle of NOMA relies on exploiting the multiplexing in the power domain to allow multiple users to communicate on the same orthogonal resources [9]. NOMA can achieve this by using superposition coding (SC) at the BS and successive interference cancellation (SIC) at the Rx side. Instead, while SC and SIC, which are the main components of NOMA, were invented more than two decades ago, NOMA scheduling is relatively new. Fig. 2.12a illustrates the difference between OMA and NOMA scheduling in the power- and orthogonal resources- (e.g., time, frequency) domains. It can be seen that NOMA permits multiple users to be served at the same orthogonal resources, while providing different amount of power to each user. In contrast, OMA assigns only one user to each radio resource.

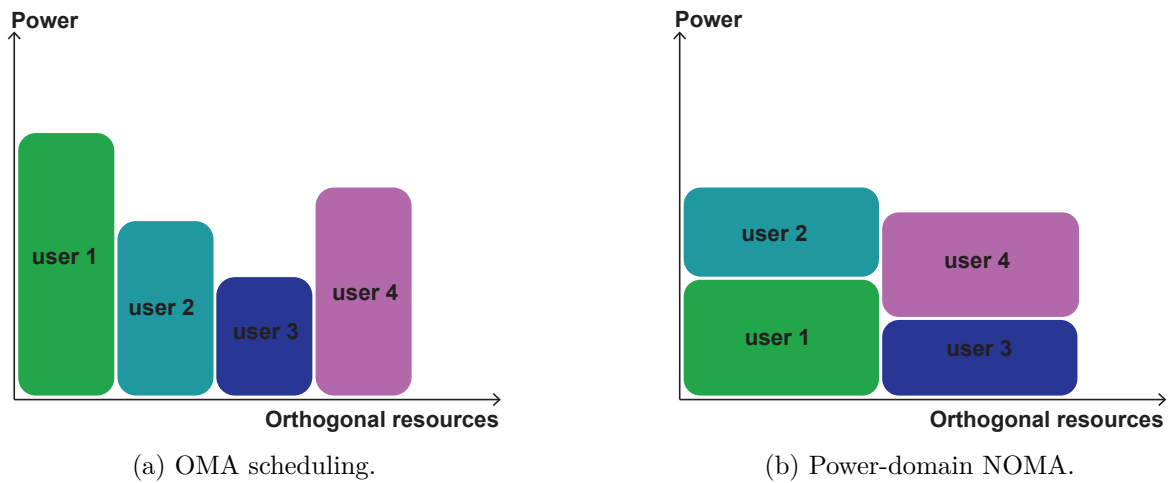


Figure 2.12: Comparison between OMA and NOMA scheduling.

Fundamentals of downlink power-domain NOMA with SIC

We now describe the general concept of NOMA downlink transmission, which consists of user multiplexing and signal separation at the transmit and receive end, respectively. We assume a downlink SISO-NOMA in a single-cell network, where a single-antenna BS simultaneously serves

K single-antenna users within the entire bandwidth B . First, the BS independently encodes and modulates the information sequence of each user. We denote s_k $k \in \{1, \dots, K\}$ the encoded and modulated symbol of user k , h_k the channel coefficient between user k and the BS, and σ_n^2 the noise power. Thus, the normalized channel gain at user k - also known as channel-to-noise ratio (CNR) - is given by $\Gamma_k = \frac{|h_k|^2}{\sigma_n^2}$.

1. Superposition coding at the transmitter side:

Using the SC process, the BS superposes the encoded and modulated symbols of the K users in the power domain and broadcasts the superimposed signal x

$$x = \sum_{k=1}^K \sqrt{p_k} s_k \quad (2.13)$$

to all users. Denote p_k as the power allocated to user k . Without loss of generality, let the K users be sorted in the descending order of their CIRs, i.e., $\Gamma_1 \geq \Gamma_2 \geq \dots \geq \Gamma_K$. Accordingly, NOMA assigns higher powers to the weak users, i.e., with lower CNRs. Thus, $p_1 \leq p_2 \leq \dots \leq p_K$, such that $\sum_{k=1}^K p_k \leq P_e$, where P_e is the total emitted power. The user k receives the following signal

$$y_k = \mathbf{h}_k \sum_{u=1}^K \sqrt{p_u} s_u + z_n, \quad (2.14)$$

where $z_n \sim \mathcal{N}(0, \sigma_n^2)$ is the received additive white Gaussian noise.

2. Successive interference cancellation at the receiver side:

Each user adopts the SIC decoding process to eliminate interference from weaker users before decoding their desired signal. However, the optimal SIC decoding order is the increasing order of CIRs. In other words, using SIC, user k first decodes the signal designated to user $k+1$ and subtracts its component from the received signal. Next, he decodes the signal designated to user $k+2$ and subtracts its component from the previously obtained signal, and so on until the signals of all users whose CIR is lower than user k are eliminated. Therefore, the user K does not perform SIC, but rather considers the signals of other users as noise. And, the user 1 decodes his desired signal without any interference if a perfect SIC is applied. Fig. 2.13 shows the case with three users served by SISO-NOMA, where we assume that the farther the users are geographically from the base station, the lower their CIRs. With a perfect SIC decoding, the received signal at user k in (2.14) can be rewritten as follows

$$y_k = \mathbf{h}_k \sum_{u=1}^{k-1} \sqrt{p_u} s_u + z_n, \quad (2.15)$$

The basic idea of NOMA is to exploit the power domain and implement multiple access with non-orthogonal resources to meet the massive connectivity requirement in 5G and beyond. In the literature, many works focus on SISO-NOMA [10,53]. Recently, MIMO-NOMA has received

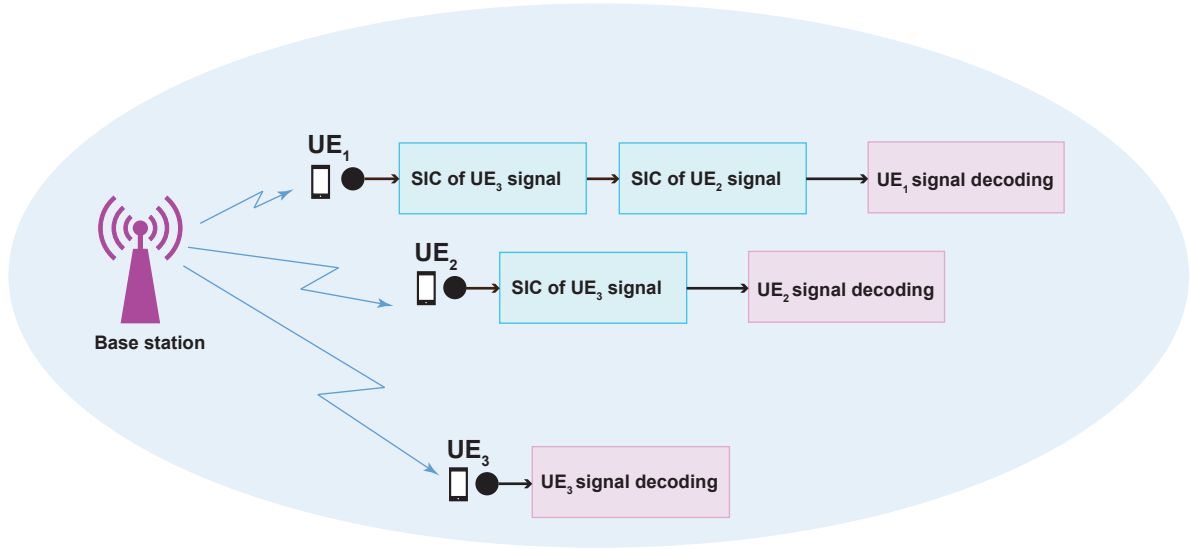


Figure 2.13: A 3-user SISO-NOMA downlink transmission with SIC decoding.

a lot of attention to improve the system capacity [54–56]. Throughout this dissertation, we focus on MIMO-NOMA. In Chapter 4, we propose a MIMO-NOMA scheme with only the angular information, in which we design user ordering, user clustering, and power allocation techniques for mmWave communications.

2.5 MmWave channels

Future radio access considers mmWave communications a promising technology because of its ultra-wide spectral bandwidth [57, 58]. The mmWave bands are generally defined as being in the 24-300 GHz frequency range, where approximately 275 GHz of total bandwidth is available. It is first necessary to study the propagation properties at these frequencies and then to design a mmWave channel model to facilitate the implementation of existing technologies at mmWave frequencies. Throughout this dissertation, we apply the 28 GHz band since it is available in the greatest number of countries across the globe and it is one of the lowest frequencies available, providing reduced device complexity and better propagation.

The wireless medium that separates Tx and Rx is called a wireless channel. Since there may be different objects, such as buildings, windows, moving cars, plants, etc., reflection, diffraction and refraction occur, and the transmitted signal propagates through different paths, as shown in Fig. 2.14. The propagation properties, including path loss, rain attenuation, atmospheric absorption, diffraction and blockage, and foliage loss behaviors, at mmWave frequencies, are very different from those at microwave frequencies. In general, the overall loss of mmWave systems is high compared to microwave systems. A channel model is a mathematical representation of how the emitted signal propagates in the wireless medium. Specifically, it includes the number of the multi-paths, the path loss, the time delay, the AoD, and the AoA of each path. Next, we give an overview of the different channel models and simulators developed in the literature.

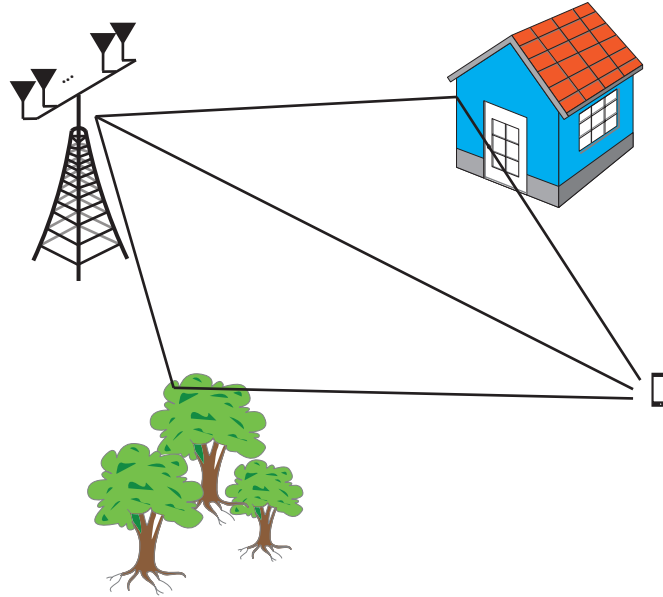


Figure 2.14: Representation of multi-path environment.

2.5.1 Review of sub-6 GHz channel models and simulators

Several channel models have been developed in the literature to facilitate the performance evaluation of wireless communication systems. The software-based models have attracted much attention since they don't require costly and complex measurements in each realization. For example, Smith [59] designs a software-based channel simulator for indoor and outdoor propagation using the multi-path Rayleigh fading channel model proposed by R. H. Clarke [60]. The Fraunhofer Heinrich Hertz institute (FHFI) designs a 3D multi-cell channel model, called QuaDRiGa [61], extended from the WINNER model to support new features, such as 3D propagation and antenna patterns, time evolution, scenario transitions, and variable terminal speeds. In [62], the authors design first a measurement-based propagation model and then a channel simulator under an indoor stair environment for IoT applications. Rappaport et al. developed statistical channel impulse response models of indoor radio channels, called SIRCIM, for both LoS and NLoS scenarios [63]. Similarly, SMRCIM (Simulation of Mobile Radio Channel Impulse Response Models) is an open-source RF propagation simulator developed for simulating outdoor channels [64]. In addition, Fung et al. developed the BERSIM channel model to simulate mobile radio communication links and calculate the average bit error rate (BER) for evaluating link quality in real-time without requiring any radio frequency hardware. WINNER II [65] is a well-known channel model for wireless networks and is developed based on a geometry-based stochastic approach to deal with the range of frequencies from 2 to 6 GHz within a bandwidth up to 100 MHz.

2.5.2 Review of mmWave channel models and simulators

Recently, the design of a mmWave channel model has attracted much attention. FHFI extends the QuaDRiGa simulator to deal with the 5G requirements and operate at mmWave frequencies. Indeed, FHFI has performed extensive channel measurements at 10/28/43/60/82 GHz. MIWEBA (Millimetre-Wave Evolution for Backhaul and Access) is a quasi-deterministic channel model developed for the mmWave outdoor mobile access links of small cell BSs with a typical cell radius of several hundred meters. The group has performed a campaign of measurements in “Leipziger Platz” in downtown Berlin, Germany. The campaign was at the 60 GHz frequency with a bandwidth of 250 MHz. However, the MIWEBA channel model requires an accurate determination of the scenario and applies to specific access use cases. The 3GPP channel model [66] is designed based on the 2D channel models from WINNER II and extended to 3D. Specifically, 3GPP is a geometry-based stochastic spatial channel model. The 3GPP channel model adopts a joint delay-angle clustering approach, and the multi-paths traveling close in both the spatial and temporal domains belong to the same cluster. In other words, a group of traveling multi-paths centered around a mean propagation delay must depart and arrive from a unique angle of departure (AoD)-angle of arrival (AoA) combination. New York University (NYU) has developed a statistical spatial mmWave channel model and simulator called NYUSIM based on extensive measurements at mmWave frequencies in New York city in different outdoor environments. Unlike the existing channel models, NYU wireless extends the 3GPP model of ultra-high frequency through the RMS angular spreads of spatial lobes and proposes the time-cluster-spatial-lobe (TCSL) clustering approach in the NYUSIM channel model. Specifically, TCSL groups the paths traveling close in time within the same time cluster (TC) and that arrive from potentially different directions in a short window of time propagation. The spatial Lobe (SL) represents the main directions of departure (or arrival), where energy arrives over several hundred nanoseconds. This TCSL approach is proposed based on the field measurements that reveal that paths belonging to different TCs can arrive at the same pointing angles, detectable due to high gain directional arrays. However, the current models, as 3GPP, have not modeled this feature, i.e., coupling the time and space dimensions. In [21], the authors compare 3GPP and NYUSIM channel models, and they find that NYUSIM is more realistic than 3GPP. Moreover, the authors in [67] compare QuaDRiGa and NYUSIM channel models, and they find that NYUSIM shares greater spatial consistency. To this end, we adopt NYUSIM, the realistic and the statistical spatial mmWave channel model, in our study.

More detailed descriptions of the mmWave channel model developed by NYU are provided below.

2.5.3 NYUSIM mmWave channel model and simulator

The NYU WIRELESS center performed extensive channel measurements from 2012 to 2017 at frequencies from 28 to 73 GHz in New York City in diverse outdoor environments, e.g., urban microcell (UMi), urban macrocell (UMa), and rural macrocell (RMA) scenarios. Based on the acquired data with a total of 1 Terabyte and the measurement analysis done in [2], NYU

WIRELESS develops statistical channel modeling software known as NYUSIM for broadband 5G and beyond wireless communications. NYUSIM can handle a large spectrum of carrier frequencies from 500 MHz to 100 GHz and support RF bandwidth from 0 Hz (continuous wave) to 800 MHz for different environment types, including UMi, UMa, and RMA. Specifically, NYUSIM is a Matlab-based open-source simulator implementing the 3D mmWave statistical spatial channel model (SSCM). Fig. 2.15 shows the graphical user interface (GUI) of NYUSIM.

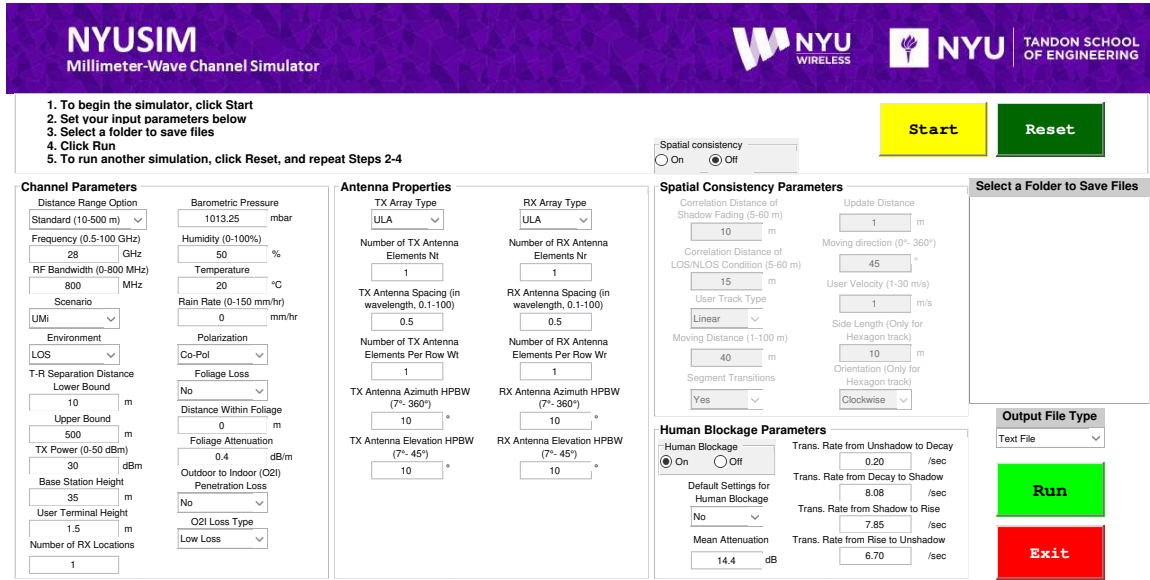


Figure 2.15: Graphical user interface of NYUSIM Version 2.

2.5.3.1 Time-cluster and spatial-lobe approach

As mentioned before, the SSCM implemented in NYUSIM utilizes the TCSL clustering technique to model the mmWave channels via TCs and SLs. TC consists of the paths traveling close in time that arrive from potentially distinct directions in a small window of time propagation. SL represents the arrival and departure directions of multi-path components (MPCs). The TCSL approach couples the temporal and the spatial channel parameters since the field observations reveal that paths belonging to different TCs can arrive at the same pointing angles. In addition, SSCM adopts a 25-ns minimum inter-cluster void interval to segment the MPCs based on their arrival times, such that consecutive MPCs arriving within a period less than 25 ns were assumed to compose one TC. The value of 25 ns is derived from the field measurements and represents the minimum propagation time between possible obstacles generating scattering, reflection, or diffraction. Fig. 2.16 depicts the omnidirectional power delay profile (PDP) and Fig. 2.17 depicts the corresponding AoD and AoA power spectrum. It can be seen that SSCM generates multiple paths and groups them in four TCs according to the propagation time. In addition, based on their AoD/AoA directions, SSCM groups the MPCs into one AoD SL and two AoA SLs.

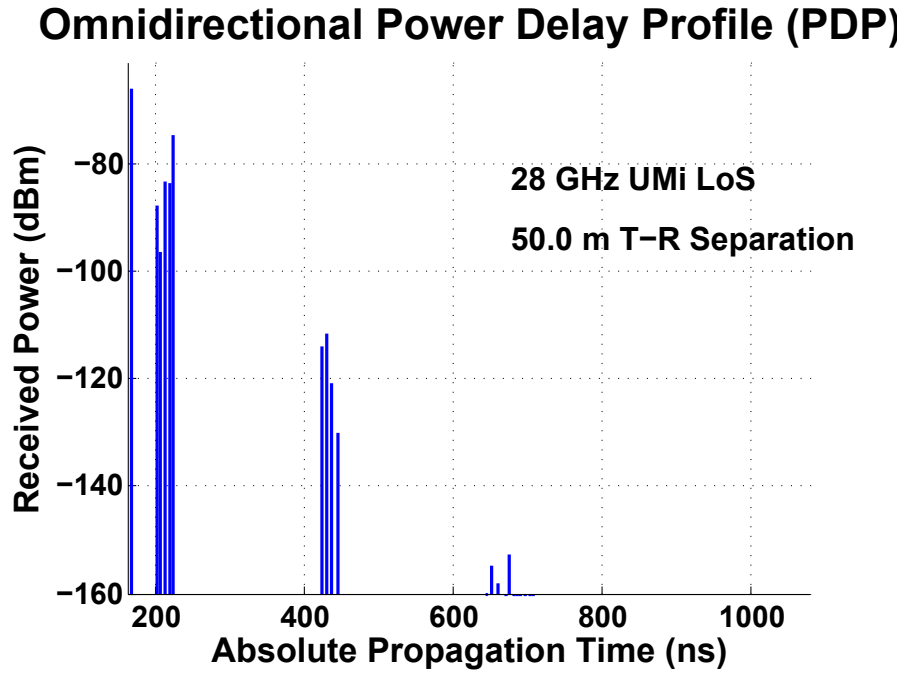
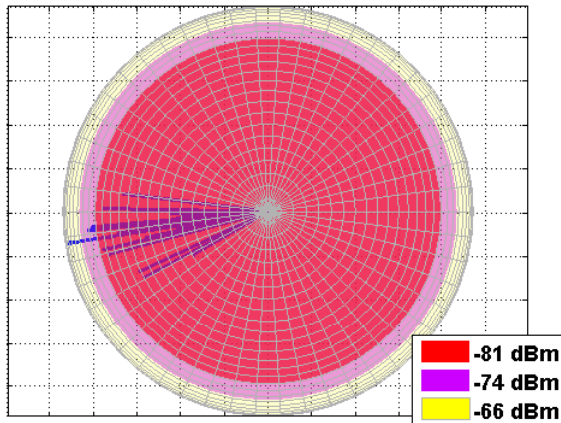


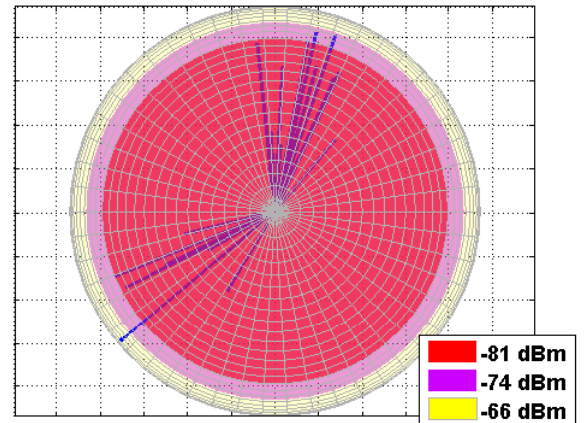
Figure 2.16: Omnidirectional power delay profile (PDP).

3-D AOD Power Spectrum - 28 GHz, UMi LoS, 50.0 m T-R Separation



(a) AoD power spectrum

3-D AOA Power Spectrum - 28 GHz, UMi LoS, 50.0 m T-R Separation



(b) AoA power spectrum

Figure 2.17: AoD and AoA power spectrum.

2.5.3.2 Input parameters

As shown from GUI in Fig. 2.1, there are 49 input parameters grouped into four main categories: channel parameters, antenna properties, spatial consistency parameters, and human blockage parameters. The panel of "channel parameters" contains the parameters describing the propagation channel, such as frequency, RF bandwidth, scenario type (which can be UMi, UMa or RMa), environment type (LoS or NLoS), lower and upper bounds of separation distance between Tx and Rx, transmit power, base station height, user equipment (UE) height, humidity, temperature, etc. The panel of "antenna properties" contains the input parameters describing

the antenna arrays at both Tx and Rx, such as array type (ULA or URA), antennas number, number of horizontal antennas, inter-element spacing distance, etc at both Tx and Rx. It is worth noting that the azimuth and elevation half-power-beamwidth (HPBW) of the Tx (resp. Rx) antenna is taken into consideration when one antenna exists at the Tx (resp. Rx). The panel of "spatial consistency" contains the input parameters describing the implementation of the spatial consistency, such as correlation distance of shadow fading, correlation distance of shadow fading, UE track type, track distance, moving direction, UE velocity, etc. The panel of "human blockage" contains the input parameters describing the human blockage shadowing loss due to a person near UT, such as the mean attenuation of blockage events, the transition rates from unshadowed state to decay state, from decay state to shadowed state, from shadowed state to rise state, from rise state to unshadowed state. In our study, both the spatial consistency and the human blockage are turned "off" since these two panels are recently added in the simulator.

2.5.3.3 Output parameters

Once the input parameters are set, NYUSIM conducts a Monte Carlo simulation to develop the CIR at each user location, assuming that channels for different user locations are independent, justified by the measurements that reveal small decorrelation distances [22, 68]. Specifically, NYUSIM generates separately the 3D mmWave temporal and spatial channel parameters. Then, it couples the obtained parameters using the TCSL approach and designs the omnidirectional/directional CIRs in both time and space domains.

The step procedure for generating the channel coefficients and the statistical distributions in all steps are detailed in [22], and can be summarized as follows

1. Generate the distance d_{TR} separating the user from the BS.
2. Generate the total omnidirectional power received at the user according to the environment type and d_{TR} .
3. Generate the number of TCs N_{TC} and the numbers of AoD and AoA SLs (L_{AoD} , L_{AoA}).
4. Generate the number of sub-paths (SPs) M_n in each TC.
5. Generate the intra-cluster SP excess delays in units of nanoseconds.
6. Generate the cluster excess delays. This step also ensures no inter-cluster overlap in time with a 25-ns minimum inter-cluster void interval.
7. Generate the power level of each TC.
8. Generate the power of each SP based on that of the TC in which belongs.
9. Generate the phase of each SP.
10. Recover the absolute time delay of each SP using d_{TR} obtained in Step 1.

11. Generate the mean AoA and AoD azimuth angles of the 3D SLs to avoid overlap of lobe angles in azimuth.
12. Generate the mean AoA and AoD elevation angles of the 3D SLs in elevation.
13. Assign a AoD and AoA SLs to each SP and generate the AoD/AoA angles OF each SP using the SL angles found in Step 11.

It is worth noting that in the NYUSIM channel model, the number of TCs varies from 1 to 6, and the number of SLs is upper-bounded by 5, and its mean is equal to 2. In contrast, the number of clusters is unrealistically large in the channel model of 3GPP TR 38.901 Release 14 for frequencies above 6 GHz [69]. For instance, the number of TCs (which is equal to the number of SLs) equals 12 and 19 in the LoS and NLoS environments, respectively, for the UMi street canyon scenario. However, the real-world measurements at mmWave frequencies do not support these impractical numbers [6, 22, 70, 71]. It can be seen that the number of clusters in the 3GPP channel model is much larger than those in the NYUSIM channel model, which are obtained according to the field measurements.

2.6 Summary

This chapter introduced a background theory of m-MIMO and NOMA and a literature review of sub-6 GHz and mmWave channel models. Throughout this dissertation, we aim to reduce the implementation complexity of the beamforming technique at the BS, and on the other hand, reduce the channel estimation overhead needed for signal processing at the BS. Moreover, we adopt the realistic and statistical mmWave channel model, called NYUSIM. Therefore, in the next chapter, we will investigate the spatial behavior and potentiality of the low-complex AD-DBF technique for mmWave m-MIMO systems using NYUSIM.

Chapter 3

Angle-Domain Digital Beamforming for MmWave Massive MIMO Systems

3.1 Introduction

As noted in Chapter 1, the demands of cellular data traffic grow faster and faster, in addition to the emergence of bandwidth-hungry applications. The spectrum below 6-GHz is already congested, and there is not enough bandwidth available to meet these requirements. Therefore, the new generation of wireless communication systems is migrating to the higher frequency bands - also known as mmWave bands - with large chunks of wide bandwidth available to achieve high data rates. Furthermore, m-MIMO can further boost the spectrum efficiency, by extensively exploiting the spatial multiplexity and diversity gains resulting from implementing massive antennas at the BS. Because the wavelength of these frequencies is tiny, a large number of antenna elements can be integrated into a small form factor. Moreover, they provide high spatial gain to overcome the severe mmWave path loss. Accordingly, the beamforming with a high-directional array is investigated as an appealing technique for mmWave wireless systems.

To get the most out of m-MIMO systems, both user and the BS should have a perfect knowledge of the CSI. In practice, they cannot know the CSI perfectly. In general, the user obtains the CSI during a training phase, in which the BS transmits training signals pre-recognized by the user. Subsequently, the BS can know the transmit-side CSI by leveraging channel reciprocity in the TDD mode [72] or via feedback channel in the FDD mode [73]. As mentioned in Section 2.3.4, the canonical form of m-MIMO operates in the TDD mode because the overhead of channel estimation scales with the number of users and not the number of antennas. Unfortunately, the reciprocity of DL and UL channels in TDD systems is challenging in practice. For instance, the inter-cell pilot contamination is the essential factor limiting the performance of TDD-based m-MIMO systems [47, 74]. Recently, some works addressed the pilot contamination problem [75–77]. Yet, in practice, other deficiencies cause channel estimation errors, which result

in non-reciprocal DL and UL channels and thereby degrade the performance of TDD systems. These deficiencies include the calibration error in the DL and UL RF chains [78], the hardware impairments at the BS and users [79, 80], etc. Besides, due to the absence of DL training for CSI estimation in TDD-based m-MIMO systems, the user cannot know the DL CSI, which may cause a significant error in data decoding applied at the user side [74]. In addition, the majority of current cellular systems deploy FDD [73, 81]. To this end, FDD-based m-MIMO systems have recently attracted a lot of attention [82–85], where the authors aim to reduce the channel estimation overhead. Throughout this thesis, we exploit the main characteristics of the mmWave channels, i.e., high directionality and blockage, that make angular information as a promising partial CSI [86], whose feedback overhead does not increase with the number of antennas. Besides, angular information that only depends on the user mobility can be adopted in TDD or FDD mode.

The precoding or BF technique is one of the essential components of signal processing to ensure the reliability of downlink transmission and improve the network capacity [87]. In the literature, different digital precoding techniques, which are classified into two main categories: 1) linear precoding and 2) non-linear precoding, have been proposed. The latter category includes dirty paper coding (DPC) [88], lattice-reduction-aided broadcast precoding [89], and vector perturbation [90] techniques. However, non-linear precoding techniques suffer from high computational complexity. With LSAAs, linear precoding techniques are more advantageous than non-linear precoding techniques due to their good performance and computational efficiency [51, 84]. To this end, the linear precoding techniques are widely considered with m-MIMO systems, such as conjugate beamforming (CB), zero-forcing (ZF), regularized zero-forcing (RZF), etc. CB maximizes the BF gain at the intended user, ZF eliminates inter-user interference, and RZF trades off the advantages of CB and ZF. ZF and RZF are more complex than CB due to the implementation of matrix inversion. In [51], the authors compare the performance of these linear precoders for m-MIMO systems in both single-cell and multi-cell scenarios.

However, the use of fully DBF with m-MIMO at mmWave frequencies poses many challenging issues, including power consumption, complexity, and channel estimation overhead. Indeed, the implementation of a large number of antennas is not as challenging as the cost and power consumption of DACs and RF chains assigned to each antenna element. In addition, the baseband of the existing DBF techniques becomes more complex with LSAAs. Besides, most of them requires full CSI, which is not easy to achieve in practice. Indeed, when m-MIMO is applied, acquiring an accurate CSI can consume excessive spectrum resources [46]. Recently, to overcome the high power consumption of the mmWave front ends, solutions as low-resolution DACs have been proposed in [91]. To reduce the overhead of channel estimation in FDD-based m-MIMO systems, many works in the literature have been proposed different training and linear precoding techniques with limited feedback [73, 82–85] by utilizing long-term channel statistics, such as spatial and temporal correlations.

In this chapter, we focus on reducing the complexity of the DL precoding and the overhead of channel estimation by leveraging the potentiality of angular information in the highly direc-

tional mmWave channels. Accordingly, throughout this chapter, we adopt AD-DBF [92, 93], the constant-envelope beamformer that uses only digital phase shifters and requires only a geometric partial CSIT, mainly the users' angular information. Unlike prior work using a simplified mmWave channel model [92, 93], we consider the realistic and statistic channel model NYUSIM from New York University. NYUSIM provides 2D and 3D MU-MIMO channel models using different antenna array architectures, namely ULA and URA. As a reference, we consider CB, which is less complex than ZF and RZF and maximizes the BF gain at the desired user, based on a full CSIT [50, 51]. The purpose of this chapter is to study the spatial behavior of AD-DBF and its potentiality for mmWave m-MIMO systems as well as providing a reference for the next chapter. The work presented in this chapter has appeared in the *24th International ITG Workshop on Smart Antennas (WSA), 2020* [94].

3.2 Channel and system model

Consider a 3D mmWave m-MIMO system, where a central BS is equipped with a 2D URA array of $M = M_x \times M_z$ antennas. M_x and M_z respectively represent the number of horizontal antennas along the x -axis and the number of vertical antennas along the z -axis; see Fig. 3.1. Both the horizontal and the vertical antennas are separated by the same inter-element spacing distance, d . In both azimuth and elevation domains, the BS concurrently communicates with K single-antenna UEs, which are randomly distributed in the cell. Moreover, we assume that the number of antennas is greater than the served users, i.e., $M > K$, generally considered in m-MIMO systems.

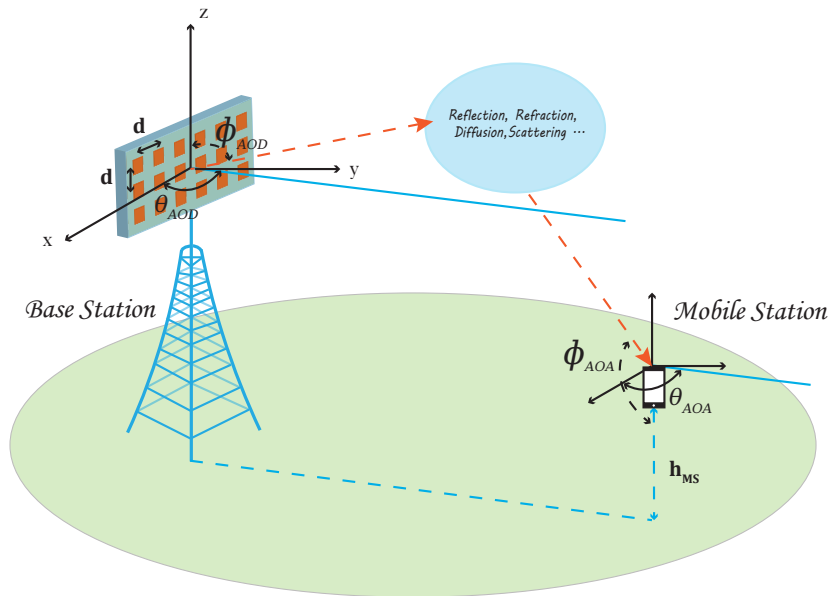


Figure 3.1: Downlink MU-MIMO system.

3.2.1 Channel model

We assume a narrowband multi-path channel for each UE, where the LoS path is not blocked. The downlink channel vector $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ between the k -th UE (denoted in this chapter as UE $_k$) and the BS is given by [95]

$$\mathbf{h}_k = \sum_{n=1}^{N_k} \alpha_{k,n} e^{j\varphi_{k,n}} \mathbf{a}^H(\vec{\Theta}_{k,n}), \quad (3.1)$$

where N_k is the number of MPCs in UE $_k$ channel, $\varphi_{k,n}$ and $\alpha_{k,n}$ are the phase and the amplitude of the n -th MPC in UE $_k$ channel, $\vec{\Theta}_{k,n} = (\theta_{k,n}, \phi_{k,n})$ is the couple of azimuth and elevation AoD of the n -th MPC in UE $_k$ channel. All these channel coefficients are generated by the NYUSIM simulator presented in Section 2.5.3. Moreover, the channel vector of each user is one of the outputs of NYUSIM. Denote $\mathbf{a}(\vec{\Theta}) \in \mathbb{C}^{M \times 1}$ as the transmit array steering vector corresponding to the $\vec{\Theta}$ direction and is given by (2.3) by setting $u = x$ and $v = z$. For simplicity, we will use, in the rest of this chapter, the notation $\mathbf{a}_{k,n}$ instead of $\mathbf{a}(\vec{\Theta}_{k,n})$. Note that the LoS path is the path with the highest power generated by NYUSIM and is labeled by $n = 1$.

3.2.2 Data transmission model

In the DL transmission, with linear precoding techniques, the signal transmitted from the M antennas, $\mathbf{x}$, is a linear combination of the modulated symbols s_k ($\forall k \in \{1, \dots, K\}$), each corresponding to one of the K UEs, as shown in Fig. 3.2. Accordingly, the mathematical expression of the transmitted signal $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is given by

$$\mathbf{x} = \sqrt{\eta} \mathbf{W} \mathbf{s} = \mathbf{W}_n \mathbf{s}, \quad (3.2)$$

where $\mathbf{s} \in \mathbb{C}^{K \times 1}$ contains the encoded and modulated symbols of all UEs, $\mathbf{W} \in \mathbb{C}^{M \times K} = [\mathbf{w}_1, \dots, \mathbf{w}_K]$ is the linear precoding matrix, with $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the precoding vector assigned to UE $_k$. We assume that the power of the emitted source signal is normalized, i.e., $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}$. Note that $\eta = \frac{1}{\text{Tr}(\mathbf{W}^H \mathbf{W})}$ is the normalization factor that eliminates the beamforming effect on the transmission power. Hence, the normalized precoding matrix $\mathbf{W}_n \in \mathbb{C}^{M \times K}$ is given by $\mathbf{W}_n = \sqrt{\eta} \mathbf{W}$. Thereby the received signal $\mathbf{y} \in \mathbb{C}^{K \times 1}$ can be expressed as follows

$$\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{n} = \mathbf{H} \mathbf{W}_n \mathbf{s} + \mathbf{z}_n \quad (3.3)$$

where $\mathbf{H} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \dots \ \mathbf{h}_K^T]^T \in \mathbb{C}^{K \times M}$ is the MU-MIMO channel matrix, and $\mathbf{n} \in \mathbb{C}^{K \times 1}$ is the additive white Gaussian noise vector with $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2)$.

3.2.3 Fully digital beamforming

In this section, we present two different DBF techniques, namely CB and AD-DBF. These beamformers differ in their purpose, complexity, and CSIT requirements.

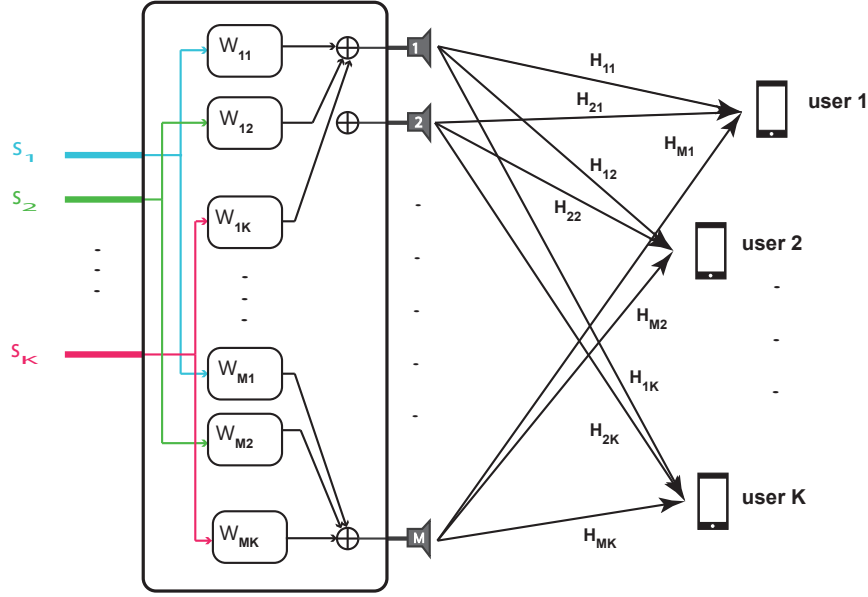


Figure 3.2: System model of downlink MU-MIMO digital beamforming.

3.2.3.1 Conjugate beamforming

CB - also known as Maximum Ratio Transmission (MRT) precoder, maximizes the BF gain at the desired user, using full CSIT. The CB beamformer is generated in terms of the MU-MIMO channel $\mathbf{H}$, and the corresponding beamforming matrix $\mathbf{W}^{CB}$ is given by [50]:

$$\mathbf{W}^{CB} = \mathbf{H}^H. \quad (3.4)$$

The beamformer envelope depends on $\mathbf{H}$. Hence, the peak-to-average power ratio (PAPR) of the waveform is directly affected. Moreover, CB requires an instantaneous full CSIT, i.e., the MU-MIMO channel matrix $\mathbf{H}$. Therefore, the rate of feedback links increases with the number of antennas and is challenging with LSAAAs. The CB aims to add constructively linear combinations of the signal. This summation grows proportionally with the number of transmit antennas, M . It should be noted that CB maximizes the signal and BF gain at each user without taking any action over the inter-user interference. And, it is the counterpart of the maximal-ratio combining (MRC) technique at the receiver side for the UL transmission.

3.2.3.2 Angle-domain digital beamforming

The AD-DBF beamformer - also known as digital beamsteering (DBS) - forms and steers a beam toward the intended user using only digital phase shifters. The AD-DBF beamformer is constructed based on the steering matrix $\mathbf{A}_{LoS} \in \mathbb{C}^{M \times K}$ corresponding to the LoS path. The DBS beamforming matrix can be expressed as follows [92]

$$\mathbf{W}^{DBS} = \mathbf{A}_{LoS}, \quad (3.5)$$

where $\mathbf{A}_{LoS} \in \mathbb{C}^{M \times K}$ is the steering matrix corresponding to the LoS path and is given by

$$\mathbf{A}_{LoS} = \begin{bmatrix} \mathbf{a}_{1,1} & \mathbf{a}_{2,1} & \cdots & \mathbf{a}_{K,1} \end{bmatrix}. \quad (3.6)$$

The precoder envelope is constant. Thus, the PAPR is minimized and the energy efficiency of the amplifier is optimized. Interestingly, DBS is a constant-envelope and a low-complex DBF technique that requires only the angular information of the LoS path of UE $_k$, i.e., $\vec{\Theta}_{k,1}$. Hence, the feedback rate is reduced when using DBS at the BS compared to that when using full CSIT-based DBF techniques, such as CB.

3.3 Sum-rate in mono- and multi-path environments

Throughout this dissertation, we adopt the sum-rate as the performance metric. In this section, we derive the system sum-rate R_T when the BS applies DBS or CB in mono- and multi-path environments.

According to the Shannon capacity formula, the achievable sum-rate R_T is given by

$$R_T = \sum_{k=1}^K R_k = \sum_{k=1}^K \log_2(1 + \text{SINR}_k), \quad (3.7)$$

where R_k and SINR_k are the rate and the SINR ratio achieved at UE $_k$, respectively. UE $_k$ achieved the following SINR

$$\text{SINR}_k = \frac{|\sqrt{\eta}(\mathbf{H}\mathbf{W})_{(k,k)}|^2}{\sum_{\substack{u=1 \\ u \neq k}}^K |\sqrt{\eta}(\mathbf{H}\mathbf{W})_{(k,u)}|^2 + \sigma_n^2}, \quad (3.8)$$

where $\mathbf{B}_{(i,j)}$ is the (i,j) -th entry of matrix $\mathbf{B}$ in the i -th row and j -th column.

3.3.1 Mono-path environment

Since one LoS path exists in a mono-path environment, we will use α_k , φ_k and $\mathbf{a}_k$ ($\forall k \in \{1, \dots, K\}$) instead of $\alpha_{k,1}$, $\varphi_{k,1}$ and $\mathbf{a}_{k,1}$. Thus, (3.1) can be rewritten by $\mathbf{h}_k = \alpha_k e^{j\varphi_k} \mathbf{a}_k^H$. Accordingly, the beamforming matrices of DBS and CB can be expressed as follows

$$\mathbf{W}^{DBS} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_K \end{bmatrix}, \quad (3.9)$$

$$\mathbf{W}^{CB} = \begin{bmatrix} \alpha_1 e^{-j\varphi_1} \mathbf{a}_1 & \cdots & \alpha_K e^{-j\varphi_K} \mathbf{a}_K \end{bmatrix}. \quad (3.10)$$

In a mono-path environment, DBS generates a directional beam toward the intended UE with uniformly distributed power among all UEs, as seen in (3.9). However, CB forms a directional beam toward each UE with different power allocation and phase shift, based on their channel gains, as seen in (3.10). In other words, the beam steered to the strongest UE (i.e., with the

strongest channel gain) carries the greatest beamforming power.

In a mono-path environment, UE_k achieves the following SINR_k^{DBS} when the BS adopts DBS

$$\text{SINR}_k^{DBS} = \frac{|\mathbf{a}_k^H \mathbf{a}_k|^2}{\sum_{\substack{u=1 \\ u \neq k}}^K |\mathbf{a}_k^H \mathbf{a}_u|^2 + \frac{\sigma_n^2}{\eta^{DBS} \alpha_k^2}}, \quad (3.11)$$

and achieves the following SINR_k^{CB} when the BS adopts CB

$$\text{SINR}_k^{CB} = \frac{\eta^{CB} \alpha_k^4 |\mathbf{a}_k^H \mathbf{a}_k|^2}{\sum_{\substack{u=1 \\ u \neq k}}^K \eta^{CB} (\alpha_k \alpha_u)^2 |\mathbf{a}_k^H \mathbf{a}_u|^2 + \sigma_n^2}, \quad (3.12a)$$

$$= \frac{|\mathbf{a}_k^H \mathbf{a}_k|^2}{\sum_{\substack{u=1 \\ u \neq k}}^K \frac{\alpha_u^2}{\alpha_k^2} |\mathbf{a}_k^H \mathbf{a}_u|^2 + \frac{\sigma_n^2}{\eta^{CB} \alpha_k^4}}. \quad (3.12b)$$

It is clear that the following factor $\Upsilon_{k,u} = \mathbf{a}_k^H \mathbf{a}_u$, which exists in the denominator of SINR_k for both DBS and CB, represents the spatial interference between UE_k and UE_u . From (3.11) and (3.12b), using either DBS or CB, interference experienced at UE_k is caused by the beams generated toward other UEs ($\text{UE}_u, \forall u \neq k$), with $\vec{\Theta}_u$ close to $\vec{\Theta}_k$. In addition, for CB, this interference depends on $\frac{\alpha_u^2}{\alpha_k^2}$, as shown in (3.12b). For instance, if UE_k has a strong channel gain and UE_u has a weak one, then $\frac{\alpha_u^2}{\alpha_k^2} < 1$, i.e., UE_u has low interference impact on UE_k . In contrast, when applying DBS, the interference term does not depend on the channel gain of UEs, as shown in (3.11), and thus the inter-user interference only depends on the proximity of $\vec{\Theta}_k$ and $\vec{\Theta}_u$, regardless of their channel gain.

3.3.2 Multi-path environment

In a multi-path environment, the beamforming matrix $\mathbf{W}^{DBS}$ of DBS is obtained using only the LoS path (labeled by $n = 1$) as follows

$$\mathbf{W}^{DBS} = \mathbf{A}_{LoS} = \begin{bmatrix} \mathbf{a}_{1,1} & \mathbf{a}_{2,1} & \cdots & \mathbf{a}_{K,1} \end{bmatrix}, \quad (3.13)$$

and, the beamforming matrix $\mathbf{W}^{CB}$ of CB given in (3.4) can be rewritten as follows

$$\mathbf{W}^{CB} = \begin{bmatrix} \underbrace{\sum_{n=1}^{N_1} \alpha_{1,n} e^{-j\varphi_{1,n}} \mathbf{a}_{1,n}}_{\text{superposition of multiple beams}} & \cdots & \underbrace{\sum_{n=1}^{N_K} \alpha_{K,n} e^{-j\varphi_{K,n}} \mathbf{a}_{K,n}}_{\text{superposition of multiple beams}} \end{bmatrix}. \quad (3.14)$$

In a multi-path environment, DBS generates a single beam toward each UE with uniform power allocation. However, CB generates multiple beams toward each UE in the directions of all paths

with different beamforming power according to their channel gains as shown in (3.14).

Thus, in a multi-path environment, UE_k achieves the following SINR_k^{DBS} when the BS applies DBS

$$\text{SINR}_k^{DBS} = \frac{\left| \alpha_{k,1} \mathbf{a}_{k,1}^H \mathbf{a}_{k,1} + \sum_{n=2}^{N_k} \alpha_{k,n} e^{j\varphi_{k,n}} \mathbf{a}_{k,n}^H \mathbf{a}_{k,1} \right|^2}{\sum_{u=1, u \neq k}^K \left| \sum_{n=1}^{N_k} \alpha_{k,n} e^{j\varphi_{k,n}} \mathbf{a}_{k,n}^H \mathbf{a}_{u,1} \right|^2 + \frac{\sigma_n^2}{\eta^{DBS}}}, \quad (3.15)$$

and achieves the following SINR_k^{CB} when the BS applies CB

$$\text{SINR}_k^{CB} = \frac{\left| \sum_{n=1}^{N_k} \alpha_{k,n}^2 \mathbf{a}_{k,n}^H \mathbf{a}_{k,n} + \sum_{n=1, n \neq m}^{N_k} \sum_{m=1}^{N_k} \alpha_{k,n} \alpha_{k,m} e^{j(\varphi_{k,n} - \varphi_{k,m})} \mathbf{a}_{k,n}^H \mathbf{a}_{k,m} \right|^2}{\sum_{u=1, u \neq k}^K \left| \sum_{n=1}^{N_k} \sum_{m=1}^{N_u} \alpha_{k,n} \alpha_{u,m} e^{j(\varphi_{k,n} - \varphi_{u,m})} \mathbf{a}_{k,n}^H \mathbf{a}_{u,m} \right|^2 + \frac{\sigma_n^2}{\eta^{CB}}}. \quad (3.16)$$

In a multi-path environment and when using DBS, the interference factor is defined as $\Upsilon_{k,u,n} = \mathbf{a}_{k,n}^H \mathbf{a}_{u,1}$, as seen in the denominator of (3.15), which represents interference of the LoS beam of UE_u ($\forall u \neq k$) (represented by $\mathbf{a}_{u,1}$) with all the paths in UE_k 's channel, (represented by $\mathbf{a}_{k,n}^H$). The inter-user interference is high if $\vec{\Theta}_{k,n}$ is close to $\vec{\Theta}_{u,1}$. Accordingly, even if all UEs are not located in the same direction, there is an interference resulting from the NLoS paths.

In the case of CB, the interference factor is defined as $\Upsilon_{k,u,n,m} = \mathbf{a}_{k,n}^H \mathbf{a}_{u,m}$. So, interference results from the collision of multiple beams $\mathbf{a}_{u,m}$ generated at each UE_u with all paths $\mathbf{a}_{k,n}^H$ of UE_k channel. It also depends on their channel gains as shown in (3.16).

Moreover, for DBS, the received power, observed at the numerator in (3.15), depends on the received power of the LoS path $\alpha_{k,1} \mathbf{a}_{k,1}^H \mathbf{a}_{k,1} = \alpha_{k,1} M$ plus the effect of paths in the vicinity of the LoS path, i.e., such that $\mathbf{a}_{k,n}^H \mathbf{a}_{k,1}$ is close to M . In contrast, for CB, the received power comes from the received power $\alpha_{k,n}^2 \mathbf{a}_{k,n}^H \mathbf{a}_{k,n}$ for each path of UE_k plus the effect of paths that are close to each other. We can conclude that, unlike CB, DBS does not have the potentiality of NLoS paths and loses some of the radiated energy.

3.4 Illustrative results and discussions

In this section, we evaluate the performance of DBS and CB in terms of sum-rate. Specifically, we investigate the impact of different factors on the DBF techniques, such as the antenna array architecture, the number of users, the number of MPCs, the antennas number, and the angle spread. We consider a perfect CSIT, i.e., perfect channel estimation at each user and instantaneous channel feedback. The temporal and spatial parameters are generated using NYUSIM. The achievable sum-rate is averaged over 2000 channel realizations. System parameters used for performance assessment are listed in Table 3.1. The noise power in Table 4.1 represents the thermal noise power at a temperature of 25°C.

Table 3.1: Simulation parameters

Parameters	Value
Carrier frequency	28 [GHz]
Channel bandwidth	20 [MHz]
Cell radius	100 [m]
Antenna separation distance	$\lambda/2$
Transmission power	30 [dBm]
Noise power	-101 [dBm]
Number of paths per TC in a rural environment [22]	$\mathcal{U}\{1, 2\}$
Number of paths per TC in an urban environment [22]	$\mathcal{U}\{1, 30\}$

3.4.1 Impact of the number of users

In this section, we assume that the BS is equipped with $M = 64$ antennas and simultaneously serves $K \leq 45 < M$ users. In addition, we consider three antenna array architectures, namely $(M_x = 64, M_z = 1)$ ULA, $(M_x = 32, M_z = 2)$ URA, and $(M_x = 16, M_z = 4)$ URA. And, we evaluate the performance of DBS and CB against the number of users, K , in an urban mono- or multi-path environment.

3.4.1.1 Mono-path environment

In Fig. 3.3, the system sum-rate with either DBS or CB is presented as a function of the number of simultaneously served users, K , in an urban mono-path environment. This figure shows clearly that DBS outperforms CB in a mono-path environment using either a ULA or a URA architecture.

To better understand this result, we study the spatial behavior of DBS and CB. The spatial distribution on the xoy plane of five UEs served simultaneously by a ULA array and their beam patterns, using DBS, are shown in Fig. 3.4 for two realizations. Moreover, it is worth noting that the beam patterns in Fig. 3.4 provide only a general view of beamwidth and steering directions, with no relationship between scale and beam power. The first realization corresponds to the case where there is interference between the users, while the users are well separated in the second realization. Moreover, the rate R_k and the channel gain α_k of UE_k with $k \in \{1, 2, 3, 4, 5\}$ are summarized in Table 3.2.

In the first realization, as seen in Fig. 3.4a, UE_1 and UE_4 are not close to any other UEs, and therefore their rates are high, as seen in Table 3.2, using either DBS or CB. For DBS, their rates are almost equal, while UE_1 has a greater rate than UE_4 for CB, since $\alpha_1 > \alpha_4$. In contrast, UE_2 is located between UE_3 and UE_5 . In fact, since UE_5 has the strongest channel gain, while UE_2 and UE_3 have weak channel gains, $R_2^{CB} (= 0.3)$ and $R_3^{CB} (= 0.8)$ are very low, while $R_5^{CB} (= 10.5)$ is very high. For DBS, the inter-user interference significantly decreases the rate of UE_5 , $R_5^{DBS} (= 4.3)$.

In the second realization, UEs are well separated. From Table 3.2, we can see that the rate of each UE is approximately the same when using DBS, whereas using CB, the higher the user

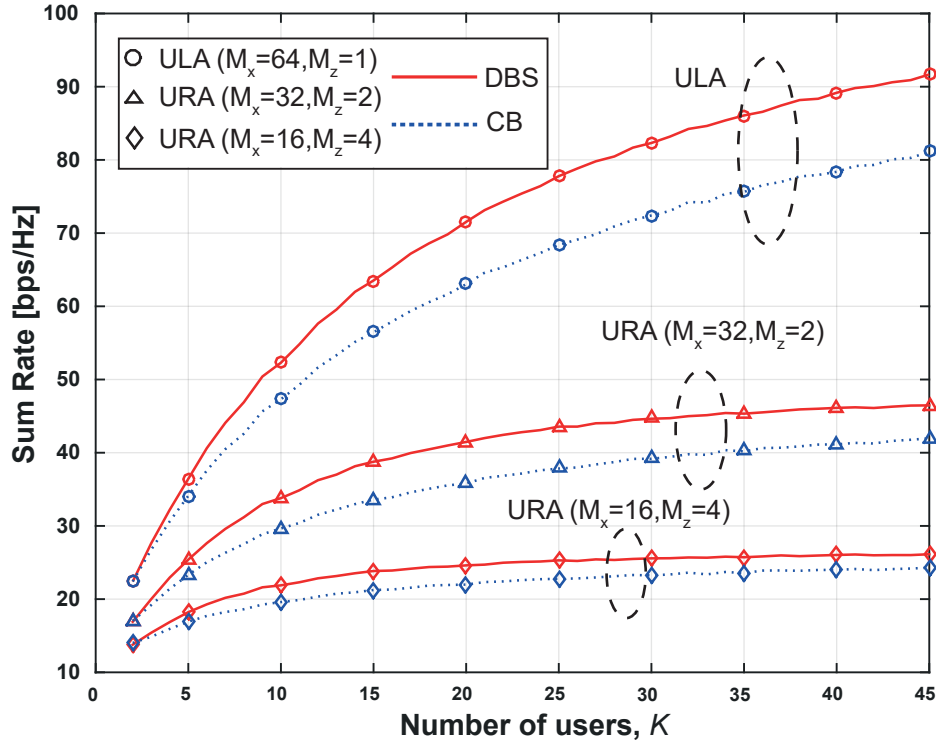


Figure 3.3: Average sum-rate achieved by a mmWave m-MU-MIMO system using either DBS or CB in an urban mono-path environment. The BS adopts a ULA or a URA array with $M = 64$ antennas.

Table 3.2: System sum-rate, rate and channel gain of each UE for two different realizations where five UEs are served simultaneously with ULA in a mono-path environment.

First realization					
Performance metrics	UE ₁	UE ₂	UE ₃	UE ₄	UE ₅
Rate R_k^{DBS}	9.9	4.2	5.8	9.6	4.3
Rate R_k^{CB}	9.8	0.3	0.8	8	10.5
Channel gain α_k	2.3e-6	1.4e-7	5.6e-8	4.8e-7	1.1e-5
Sum-rate $R_{DBS} = 33.8$, $R_{CB} = 29.4$					
Second realization					
Rate R_k^{DBS}	10.3	9.5	9.2	9.3	9.6
Rate R_k^{CB}	10.1	7.5	8.4	10.7	10
Channel gain α_k	3.5e-7	9.8e-8	3 e-7	7.5e-7	3.5e-7
Sum-rate $R_{DBS} = 47.9$, $R_{CB} = 46.7$					

channel gain is, the greater the achievable rate is. In other words, the CB performs a kind of power allocation to maximize the sum-rate without being fair to each user.

In summary, these two realizations illustrate the fact that DBS gives a moderate rate for each UE regardless of its channel gain and that CB strengthens strong UEs and marginalizes weak UEs, as already explained in Section 3.3.1.

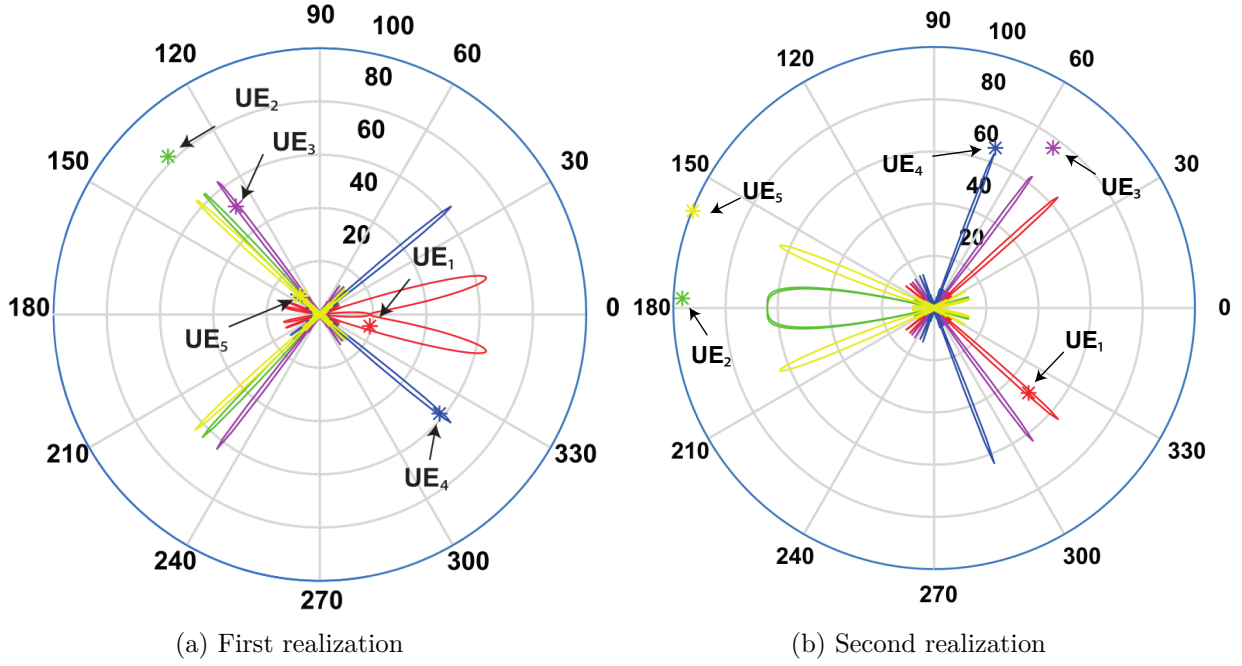


Figure 3.4: Azimuth representation of the spatial distribution of UEs (*), on the xoy plane, with their ULA beam patterns, using DBS for the two realizations in a mono-path environment. Each UE is represented by a colored * and its corresponding beam has the same color. Note that the scale corresponds to the separation distance between UEs and BS in meter, where BS is located at the center of the circle.

3.4.1.2 Multi-path environment

Due to the narrowband channel assumption, all paths belong to the same TC. In an urban multi-path environment, the number of paths N_k in UE_k channel follows a uniform distribution ranging from 1 to 30 as listed in Table 3.1.

In Fig. 3.5, the system sum-rate with either DBS or CB is presented as a function of the number of simultaneously served users, K , in an urban multi-path environment, with different antenna array architecture. We can see that the sum-rate R_T^{CB} is always larger than R_T^{DBS} when using URA. However, when the BS adopts ULA, DBS is better than CB as long as $K < 15$, and CB becomes better after that.

To understand the obtained results, we consider the case where seven UEs are simultaneously served by the BS adopting once ULA and another time 16×4 URA. For this realization and for each UE, the spatial distributions of paths, on the xoy plane, are represented in Fig. 3.6. The LoS path of each UE indicates the direction of the corresponding UE. This is corroborated by the spatial distributions of the seven UEs, on the xoy plane, represented in Fig. 3.7, along with their beam patterns obtained by applying ULA and DBS. In Fig. 3.7, the same note, about scale and beam pattern as in Fig. 3.4, is taken into account. The achievable rate of each UE is summarized in Table 3.3. Note that the channel realization of the seven users is the same with either ULA or URA.

First, as shown in Fig. 3.7, UEs are well separated in space. Therefore, since DBS generates

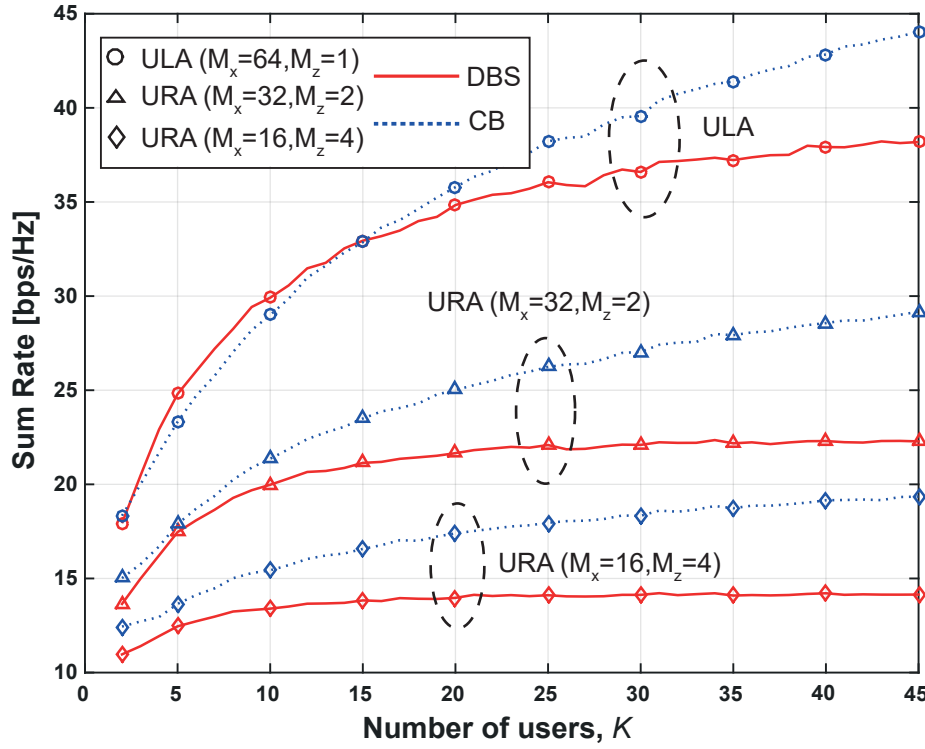


Figure 3.5: Average sum-rate achieved by a mmWave m-MU-MIMO system using either DBS or CB in an urban multi-path environment. The BS adopts a ULA or a URA array with $M = 64$ antennas.

Table 3.3: Sum-rate and rate of each UE where seven UEs are served simultaneously by ULA or 16×4 URA and using either DBS or CB. These obtained values correspond to for the same realization

Individual rate	ULA- R_k^{DBS}	ULA- R_k^{CB}	URA- R_k^{DBS}	URA- R_k^{CB}
UE ₁	6.4	2.5	1.5	1.9
UE ₂	4.1	2.6	3.1	2.4
UE ₃	5.5	3.7	0.4	2.6
UE ₄	6	2.8	2.1	2
UE ₅	9	5.7	5.3	4.8
UE ₆	1.8	1.8	0.2	0.5
UE ₇	7.5	8.7	2.5	7.5
Sum-rate	40.3	27.8	15.1	19.1

a single beam toward each UE, these beams are also well separated. Second, when applying ULA with 64 antennas, the generated beams are very narrow, as shown in Fig. 3.7. Thanks to these narrow beams, on Fig. 3.6, we observe that, except for UE₆, the NLoS paths of each user weakly interfere with the other UE beams that are in the direction of user LoS path. Thus, with ULA, DBS provides interesting sum-rate as seen in Table 3.3, with R_T^{DBS} ($= 40.3$) larger than R_T^{CB} ($= 27.8$). Similar results are shown in Fig. 3.5, where, DBS is better than CB as long as $K < 15$ with ULA.

Let's now observe the performance of DBS when applying URA instead of ULA, in Fig. 3.5.

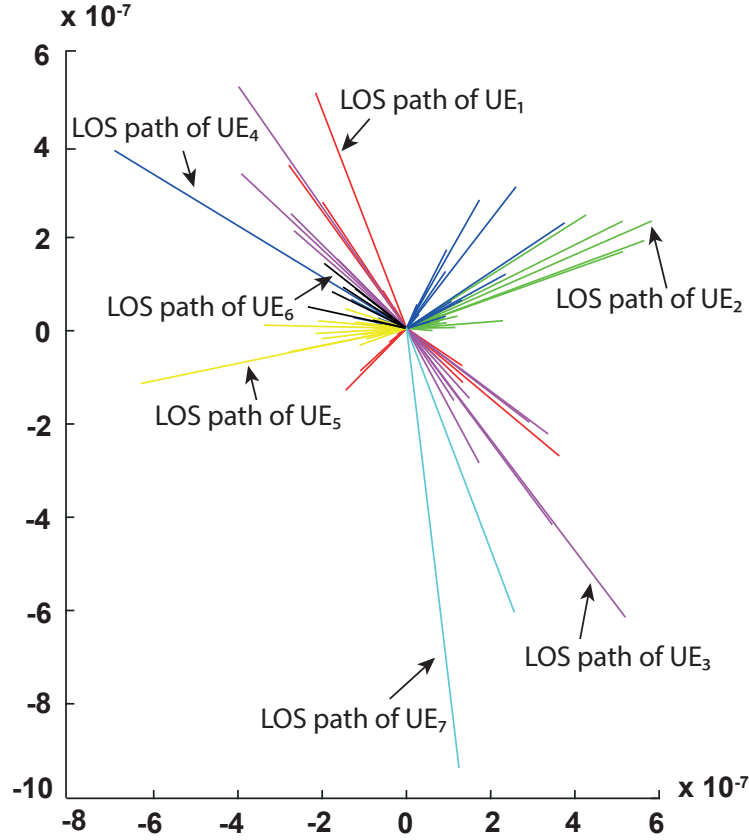


Figure 3.6: Azimuth representation of the spatial distribution of paths, on the xoy plane, for one realization in an urban multi-path environment. All paths belonging to the same UE channel are identified by the same color. The used scale corresponds to their channel gain.

ULA provides the best performance since its azimuthal resolution is able to discriminate the users spatially. With URA, the number of horizontal antennas and so its azimuthal resolution decreases, creating more inter-user interference. For ULA, thanks to a narrow beam, the NLoS of UE₁ (red) do not interfere with the beam of UE₃ (magenta). However, for URA, with wider beams, interference increases and we observe a significant degradation of the rate from R_1^{DBS} ($=6.4$), for ULA, to R_1^{DBS} ($=1.5$), for URA, as seen in Table 3.3. The same explication holds for the significant degradation of rate for UE₃, UE₄ and UE₇, when using DBS and passing from ULA to URA.

Obviously, there is also a degradation of the CB sum-rate when passing from ULA to URA due to wider beamwidth. However, this degradation is less than that of DBS sum-rate, and thus CB outperforms DBS for URA, as illustrated in Fig. 3.5. In Table 3.3, R_T^{DBS} ($=15.1$) is less than R_T^{CB} ($=19.1$) when applying a URA array with $M_x = 16$. This is mainly because DBS misses the potentiality of NLoS paths, while CB exploits their spatial diversity, as explained in Section 3.3.2.

From Fig. 3.5, we can see that the sum-rate of DBS is flat from a certain value of K , depending on the number of horizontal antennas, i.e., the azimuth beamwidth. However, the sum-rate of CB still increases when more UEs are served simultaneously. Indeed, for DBS, first, if more UEs are served by BS, then more paths exist, and interference coming from those paths

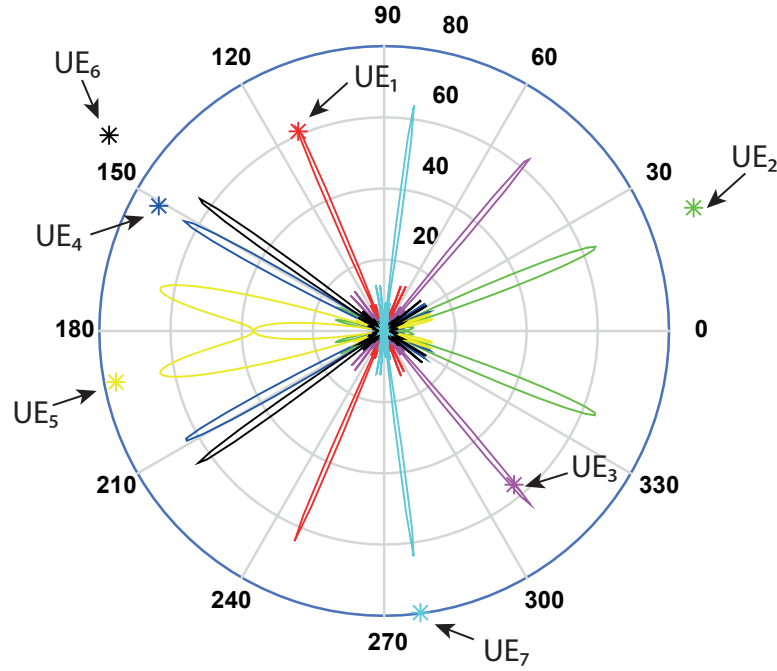


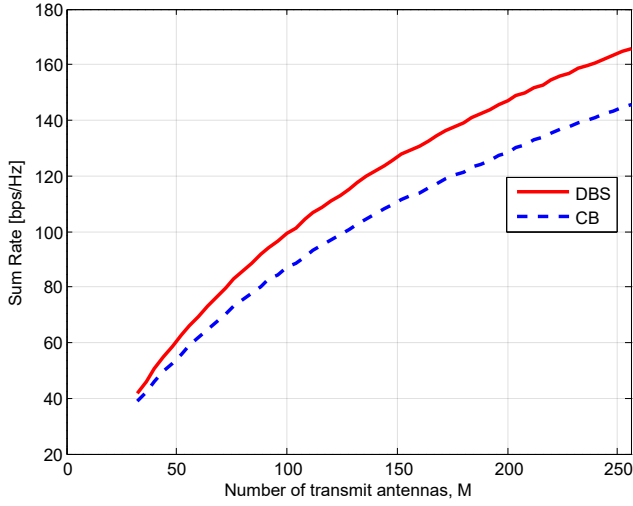
Figure 3.7: Azimuth representation of the spatial distribution of UEs (*), on the xoy plane, with their ULA beam patterns, using DBS, for one realization. Each UE is illustrated by a colored * and its corresponding beam has the same color.

(NLoS or LoS) increases. Second, DBS lacks the NLoS paths potentiality in terms of radiated energy.

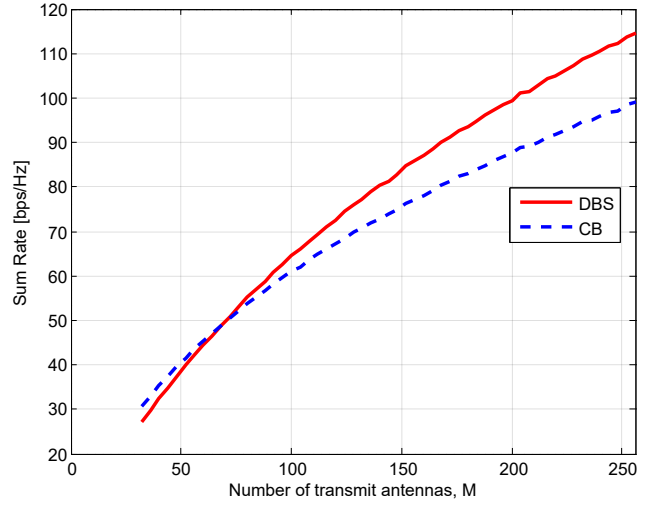
3.4.2 Impact of the number of transmit antennas

In this subsection, we investigate the impact of the number of transmit antennas, M , on the performance of DBS and CB in both rural and urban environments. In Fig. 3.8 and Fig. 3.9, the BS is equipped with ULA along the x -axis and $M_x \times M_z$ -URA, respectively, to simultaneously serve K users. Here, we set $K = 32$, and we vary the number of antennas at the BS such that $32 \leq M \leq 252$. In the case of URA, we fix the number of vertical antennas $M_z = 8$, and we vary the number of horizontal antennas, i.e., $4 \leq M_x \leq 32$. Fig. 3.8 and Fig. 3.9 illustrate the sum-rate achieved using either DBS or CB in both rural and urban environments as a function of the number of antennas, M , and number of horizontal antennas, M_x , respectively.

Fig. 3.8a and Fig. 3.9a show that DBS achieves a better performance than CB in rural environments, regardless of the number of antennas using either ULA or URA. Indeed, according to NYUSIM, one LoS path exists solely or with another NLoS path in rural channels as listed in Table 3.1. For that, the same analysis can be conducted in rural channels as in mono-path channels; see Subsection 3.4.1.1. Besides, it can be seen that the performance gap between DBS and CB increases when the number of antennas increases. When the BS adopts a large number

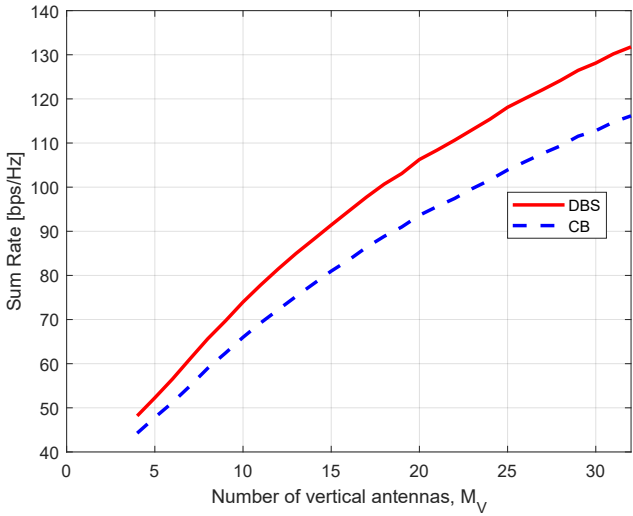


(a) Rural environment.

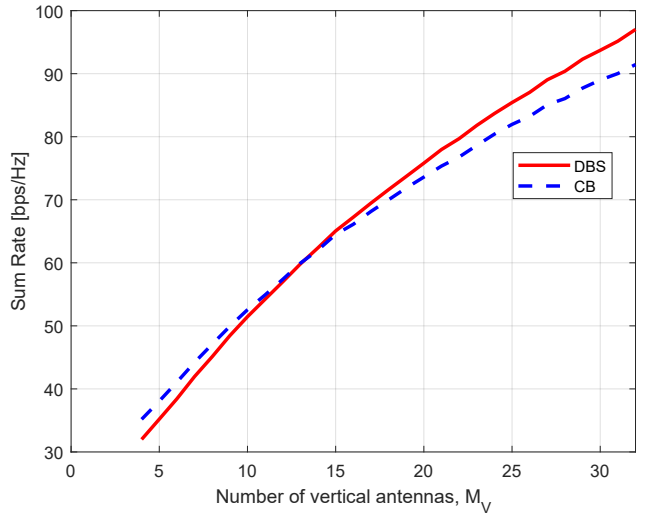


(b) Urban environment.

Figure 3.8: Average sum-rate achieved using either DBS or CB against the number of antennas, M , in a rural or urban multi-path environment when the BS is equipped with ULA and serves $K = 32$ users.



(a) Rural environment.



(b) Urban environment.

Figure 3.9: Average sum-rate achieved using either DBS or CB against the number of vertical antennas, M_v , in a rural or urban multi-path environment when the BS is equipped with URA and serves $K = 32$ users.

of antennas relative to that of users served, the narrow beam focuses the energy in the direction of the desired user, increasing the received power and reducing interference from neighboring users.

Fig. 3.8b and Fig. 3.9b show that CB outperforms DBS in urban environments when $M \leq 70$ (resp. $M \leq 8 \times 12 = 96$) using ULA (resp. URA). This is because CB exploits the potentiality of the NLoS paths in contrast to DBS, as explained in Subsection 3.4.1.2. However, when the

number of antennas increases beyond these values, DBS outperforms CB, and the gap between the curves increases with the antennas number, due to the narrow beams. In other words, with a number of antennas much greater than that of users, DBS is able to discriminate users well in azimuth, thus increasing the BF gain at each user and reducing inter-user interference. It is worth noting that when the number of antennas is much greater than that of users in a rural or urban environment, the ignorance of the NLoS paths using DBS results in a better performance of DBS against CB that considers these paths. This implies that DBS is more suitable to the m-MIMO systems with better performance, computational and estimation overhead efficiencies.

3.4.3 Impact of the angle spread

In this subsection, we consider a BS equipped with a ULA array with $M = 64$ elements along the x -axis and serving simultaneously $K \in \{16, 32\}$ users. We assume a mmWave channel with only two paths; one is LoS, and the other is NLoS. And, we vary the angle spread $\Delta\theta$ from 0° to 20° , which is defined here as the angular distance between the LoS and NLoS paths.

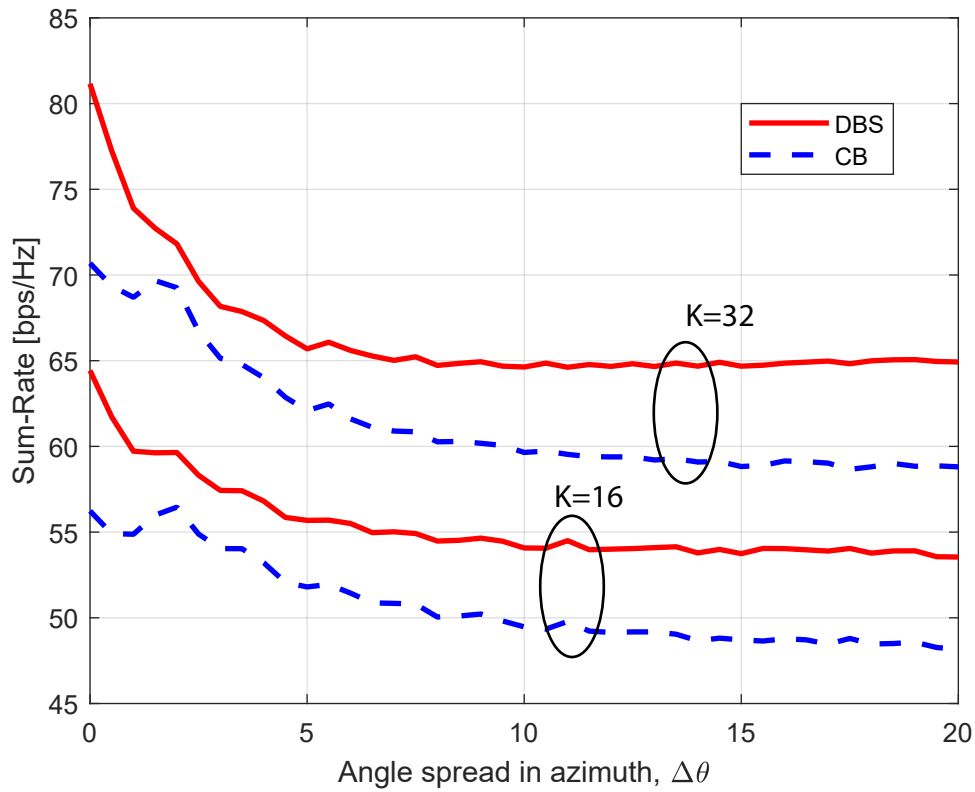


Figure 3.10: Impact of the angle spread on the average sum-rate in a two-path channel when a BS equipped with 64 linear antennas along the x -axis serves $K \in \{16, 32\}$. The sum-rate is averaged over 5000 channel realizations.

Fig. 3.10 illustrates the behavior of DBS and CB in response to varying the angular spread between the LoS and NLoS paths. As explained before, DBS outperforms CB because only two paths exist with quite narrow beams since $M = 64$ and $K \ll M$. From Fig. 3.10, we see that

the performance of both beamformers first monotonically degrades, then slowly decreases with increasing angular spread $\Delta\theta$. Accordingly, both beamformers are suitable for the highly LoS mmWave environments, i.e., when the angle spread is very small. This is because they efficiently maximize the BF gain at the intended user with a low interference coming from the NLoS path, since the MPCs belonging to the same user channel depart from close directions.

3.5 Summary

In this chapter, we considered the mmWave m-MIMO systems, and we investigated AD-DBF or DBS, the low-complex DBF technique requiring the angular information, whose feedback overhead does not scale with the number of antennas. Unlike prior works, our evaluation is done using a realistic channel model generated by NYUSIM able to provide path angle information in both rural and urban scenarios. We first derived the sum-rate for DBS and CB in mono- and multi-path environments. Therefore, while relying on simulations based on geometric aspects, we have obtained the following results

- In mono-path environments, DBS gives a moderate rate for each user, while CB reinforces the strong users and marginalizes the weak ones.
- In multi-path environments, for DBS, even when users are well separated in the space, there is an interference coming from the NLoS paths. Besides, CB maximizes the BF gain by exploiting the diversity of NLoS paths whereas DBS lacks their potentiality and loses some of the radiated energy.
- The number of users, that DBS can serve simultaneously in multi-path environments, is constrained by the number of horizontal antennas, i.e., the azimuth beamwidth.
- With a large number of antennas compared to the number of users, DBS always outperforms CB even in a multi-path environment.
- AD-DBF has a great performance in a highly LoS environment, where the paths depart from close directions.

Finally, DBS is an attractive beamformer for mmWave m-MIMO systems in terms of performance, complexity, and channel estimation overhead. However, in practice, due to the high complexity and cost of the hardware, the number of antennas cannot approach infinity, so the BS cannot completely eliminate interference between users. With a finite number of antennas, especially in a congested cell, the DBS performance degrades due to the high inter-user interference caused by neighboring beams. To solve this issue, we will propose, in the next chapter, an interference management scheme without deploying additional antennas.

Chapter 4

Multi-User Angle-Domain MIMO-NOMA for mmWave Communications

4.1 Introduction

As mentioned in Chapter 1, in mmWave wireless communications, DBF with high directional arrays is investigated as an attractive technique to mitigate the tremendous path loss and achieve high data rates [87, 96]. Interestingly, Chapter 3 reveals that AD-DBF or DBS is one of the promising DBF technique in mmWave multi-user systems in terms of sum-rate performance, complexity efficiency, and channel estimation overhead [94, 97]. Unfortunately, in a congested cell, the MU-MIMO system suffers from high inter-user interference caused by adjacent beams directed to users located in the same direction, which degrades the performance of AD-DBF. One possible solution is implementing massive antennas at the BS. Indeed, when the number of antennas approaches infinity, the multi-user interference is reduced to zero [98]. However, increasing the number of antennas is difficult in practice due to the power consumption and transceiver complexity. Recently, as mentioned in Chapter 1, NOMA has been investigated with mmWave m-MIMO systems to address the massive connectivity and further increase the spectrum efficiency exploiting the multiplexing in the power domain. As a result, coupling DBF and NOMA can help manage high user density in a congested cell by providing spatial and non-orthogonally DoFs with tens of transmit antennas (e.g., 32, 64). The case of hundreds of antennas is dealt in the next chapter.

In this chapter, we aim to design an AD MIMO-NOMA scheme for mmWave multi-user systems to reduce inter-user interference and improve the AD-DBF performance without adding additional antennas. In particular, we exploit the main features of mmWave channels, and we propose new angular-based metrics to design and develop effective user clustering and ordering and power allocation techniques. These techniques only require angular information and can be used for ULA and URA architectures. The work presented in this chapter has appeared in

the 16th International Conference on Wireless Communications & Mobile Computing (IWCMC 2020) [99], and an extended journal version has been published in IEEE Access [100]. Next, we give an overview of the related works on MIMO-NOMA.

4.1.1 Related works

NOMA-based MU-MIMO systems have been extensively studied in the literature. Most of the existing works on MIMO-NOMA design the BF weight vector, user clustering and power allocation techniques based on the assumption of perfect or imperfect full instantaneous CSIT [55, 101–109]. For instance, the authors in [101] proposed a user selection algorithm and a power schedule technique that reduce interference and enhance the performance of the ZF-based MIMO-NOMA system. A joint beamforming and power allocation scheme was proposed in [102] to maximize the sum-rate of a 2-user mmWave MIMO-NOMA system using an ABF structure with a phased array. The authors in [103] developed a user pairing and pair scheduling algorithm that improves the SE of a m-MIMO-NOMA system serving multi-antenna users. A hybrid user pairing scheme is proposed in [104] to jointly maximize the spectral and energy efficiencies of MIMO-NOMA, by applying the simultaneous wireless information and power transfer (SWIPT). In [105], the authors design m-MIMO-NOMA systems with imperfect SIC decoding. Particularly, a dynamic power allocation was proposed to maximize the achievable data rates of the weak users having worst channel conditions within each NOMA group. The authors in [106] proposed a joint user grouping and multi-dimensional resource allocation for the multi-cell DL MIMO-NOMA systems. Specifically, the authors eliminate inter-group interference by applying the ZF BF technique. In [107], the authors designed an energy-efficient power allocation method for MIMO-NOMA systems with multiple users in a cluster. A spectrum and energy efficient mmWave transmission scheme that integrates NOMA with beamspace MIMO was first proposed in [108], in order to reduce the hardware complexity and energy consumption by using lens antennas. In addition, the authors designed a precoding scheme based on the principle of ZF to reduce inter-cluster interference. For multicast MIMO-NOMA, an iterative algorithm was proposed in [109] to find the BF vector and the power for each UE that solves the power minimization problem.

However, acquiring full CSI (FCSI) is challenging due to user mobility, and on the other hand, because of the overall system overhead. One possible solution is to use statistical CSI (SCSI), such as the spatial correlation or the channel mean, which can be easily and accurately obtained at the transmitter [110]. For instance, the authors in [54] formulate an ergodic capacity maximization problem under SCSI to design the power allocation for a Rayleigh fading 2-user MIMO-NOMA system. In [111], an optimal statistical precoder for MIMO-NOMA is designed to maximize the ergodic capacity. In [112], a joint rate and power allocation scheme is proposed for downlink NOMA to minimize the total transmission power under SCSI. Furthermore, the authors in [113] modeled a partial feedback channel (e.g., quantized channel) using the random vector quantization technique. And, the proposed scheme, including the user clustering algorithm developed based on proportional fairness and the power allocation strategy that ensures the

weak users' sum capacity, is investigated with both full and quantized CSI.

However, in the context of mmWave transmissions, the FCSI challenge can be mitigated thanks to the specific properties of mmWave channels [114,115]: i) the highly directional propagation of mmWave wireless links and ii) the sensitivity of the mmWave signals to blockages. Thus, the NLoS components are severely attenuated compared to the LoS link. As a result, the spatial location of the user, i.e., the angle and the distance w.r.t. the BS, can efficiently characterize the mmWave channels [116]. Motivated by these properties, several works design NOMA systems with limited feedback for the mmWave networks. To reduce the system overhead, the authors in [86] propose a NOMA scheme with random ABF, where only the users' distance information is fed back to the BS instead of FCSI. Specifically, the BS randomly generates a single beam, and users falling in this beam are served using NOMA with a fixed power allocation strategy. Meanwhile, the users are ordered according to their distance information instead of their effective channel gain. In [117], a joint user selection and ABF scheme based on the user position is proposed for mmWave NOMA systems. First, the authors define the users' pairing angle based on the angle difference between the two users. Thereby, they develop a user selection method that assigns a near user with a far one w.r.t. the BS, which are expected to be within the specified beamwidth and have the minimum pairing angle. In [116], a location-based low-complexity user pairing algorithm is proposed for the mmWave MIMO-NOMA transmission by applying the maximum ratio transmission technique. The authors consider the Rician fading channels and derive the switching boundary between NOMA and SDMA in the angle-distance plane for the 2-user scenario. They found that NOMA is preferable when the angle difference between the two users is smaller than a threshold, or their distance difference is larger than a threshold. By applying these results in the multi-user scenario, they design a low-complexity location-based user pairing algorithm. In [118], the authors propose a user pairing strategy based on both angle and distance. The BS pairs the nearest user with another having the minimum relative angle distance, such that the latter is located within a distance threshold from the BS.

4.1.2 Contributions

Consequently, both user angle and distance information w.r.t. the BS has significant potential in the field of mmWave MIMO-NOMA systems with limited feedback. In the light of these articles, we consider a downlink MIMO-NOMA scheme with limited feedback that exploits only the ADI, i.e., the user's spatial angle¹ that are collected at the BS, to provide: 1) user clustering, 2) user ordering, 3) power allocation, 4) AD-DBF performing simultaneous addressing of clusters. In our work, ADI is assumed to be perfectly estimated and known at the BS. Figure 4.1 illustrates our proposed MIMO-NOMA scheme: thanks to the definition of an inter-user spatial interference metric, users are grouped within clusters, themselves addressed simultaneously by the beams of the digital beamformer. NOMA is used in each cluster to manage multiple access interference.

¹The user's spatial angle w.r.t. the BS represents the AoD of the LoS path and is a slowly time-varying CSI compared to the other CSI components. This information can be estimated and tracked with standard AoDs estimation algorithms such as [119,120] at the BS or the receiver side. If the users estimate this information, then they send it back to the BS.

In our work, the generalization ability is investigated. We will study not only the case of two-user NOMA groups, but also the case of multi-user NOMA groups. Besides, we consider both 2D and 3D AD MIMO-NOMA using ULA and URA, respectively.

The main contributions of this chapter can be summarized as follows

- We exploit the specific properties of mmWave channels: i) the highly directional propagation of mmWave wireless links and ii) the sensitivity of the mmWave signals to blockages. And, we derive ADI-based performance metrics, such as the user channel quality, the sum-throughput, and inter-user spatial interference, only based on ADI.
- We propose a new user ordering strategy for the NOMA protocol, with only the ADI feedback thanks to the defined user channel quality metric that considers inter-cluster spatial interference and only requires ADI.
- We derive an ADI-based inter-user spatial interference metric assuming LoS link dominance. Unlike previous works on angle-based NOMA systems that use only the spatial angular distance between users - also known as the azimuth angle difference, this metric includes the main lobe beamwidth information and does not require additional feedback information. The beamwidth is highly dependent on the number of antennas and the beam steering angle, as discussed in Section 2.5. According to this metric, we develop 2-user and multi-user clustering algorithms to reduce inter-user interference.
- A power optimization problem that maximizes the total system throughput is formulated with FCSI or ADI feedbacks and under total power and efficient SIC constraints. The obtained problem is simplified to a convex problem, thanks to a fixed inter-cluster power allocation strategy. Then, this simplification allows us to derive the closed-form intra-group power allocation solution, obtained via the Karush-Kuhn-Tucker (KKT) conditions.
- Extensive simulations are performed in mmWave multi-path rural environments to verify the performance of the proposed schemes, using NYUSIM, when the BS applies ULA. Specifically, we compare the performance of the proposed ordering strategy with other limited feedback ordering strategies. We show the effectiveness of SIC when applying the angular-based power allocation. Moreover, we compare the proposed 2-user and multi-user AD MIMO-NOMA schemes with the conventional AD-DBF (C-AD-DBF) scheme presented in Chapter 3, and the OMA scheme. In addition, we provide simulation results when the BS applies URA to verify the importance of the AD MIMO-NOMA schemes that can improve AD-DBF performance without adding additional antennas, and offers affordable hardware complexity and power consumption.

The rest of this chapter is organized as follows. Section 4.2 presents the system model of AD MIMO-NOMA that integrates AD-DBF with NOMA. In Section 4.3, we derive the ADI-based performance metrics. Section 4.4 discusses the 2-user and the multi-user clustering algorithms. In Section 4.5, we design a sub-optimal power allocation policy using either FCSI or ADI.

Section 4.6 numerically evaluates the performance of the proposed schemes, and Section 4.7 gives a summary of this chapter.

4.2 System description and problem statement

We consider a mmWave AD MIMO-NOMA system in the downlink, as depicted in Fig. 4.1, with one BS and K single-antenna UEs. The BS is equipped with a 2D array of $M = M_x \times M_z$ antennas, where M_x and M_z are the number of horizontal and vertical antennas, respectively. All UEs are represented by the ID set $\mathcal{K} = \{1, \dots, K\}$. According to the user clustering algorithm described in Section 4.4, the K UEs are grouped into $C \leq K$ spatial clusters with K_c UEs in the c -th cluster, $c \in C = \{1, \dots, C\}$. NOMA is adopted within the c -th cluster where multiple UEs exist. Using digital phase shifters and knowledge of the ADI, i.e., the user's spatial direction, AD-DBF forms a beam to each cluster.

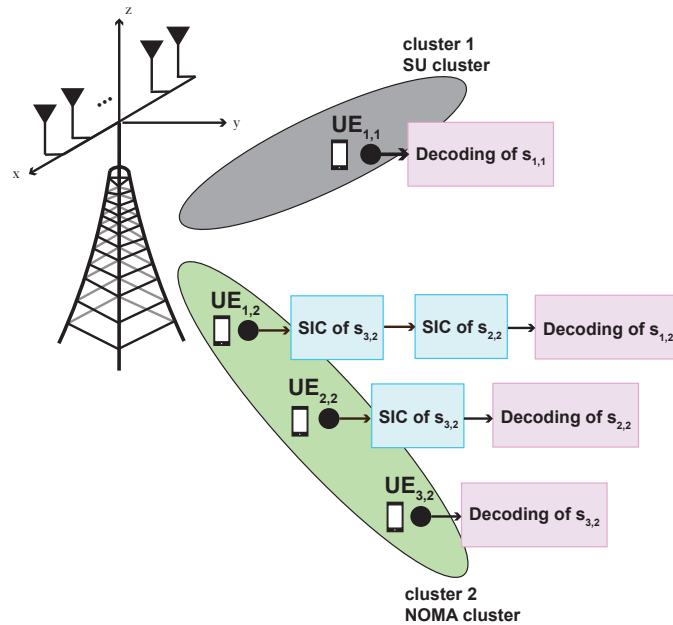


Figure 4.1: Multi-user MIMO-NOMA system.

Consequently, two different types of clusters exist, namely single-user (SU) and NOMA clusters, as shown in Fig. 4.1. In the former, only one UE is served thanks to a beam generated by AD-DBF. However, in the latter, multiple UEs share the same spatial and time-frequency resource blocks using NOMA. $C_n = \{1, \dots, C_n\}$ and $C_s = \{C_n + 1, \dots, C\}$ are the ID sets of NOMA and SU clusters, respectively, with C_n and C_s being respectively the number of NOMA and SU clusters. Besides, $\mathcal{K}_c = \{1, \dots, K_c\}$ is the newly defined ID set of UEs in the c -th cluster, with K_c being the number of UEs in the c -th cluster. Hence, $K_c = 1, \forall c \in C_s$ and $C = C_n \cup C_s$. In this work, we propose two schemes, namely 2-user and multi-user AD MIMO-NOMA. In the former, only two users exist in the NOMA cluster, i.e., $K_c = 2, \forall c \in C_n$. Instead, the latter is extended to allow for multi-user communication per NOMA cluster, i.e., $K_c \geq 2, \forall c \in C_n$.

4.2.1 Channel Model

The mmWave channel vector $\mathbf{h}_{l,c} \in \mathbb{C}^{1 \times M}$ between the BS and the l th UE in the c -th cluster, denoted in this chapter as $\text{UE}_{l,c}$, is

$$\mathbf{h}_{l,c} = \sum_{q=1}^{N_{l,c}} \alpha_{q,l,c} \mathbf{a}_{q,l,c}^H, \quad (4.1)$$

where $N_{l,c}$ is the number of MPCs in $\text{UE}_{l,c}$'s channel, $\alpha_{q,l,c}$ and $\theta_{q,l,c}$ are respectively the complex gain and the spatial angle associated with the q -th path in $\text{UE}_{l,c}$'s channel, and $\mathbf{a}_{q,l,c} = \mathbf{a}(\vec{\Theta}_{q,l,c}) \in \mathbb{C}^{M \times 1}$ is the array steering vector corresponding to $\vec{\Theta}_{q,l,c}$. In this work, we assume that the LoS path exists in the channel of each user. And, the case where the LoS path does not exist for some users is left as a future work. Denote $q = 1$ as the index of the LoS path, i.e., the strongest path. Hence, $\vec{\Theta}_{1,l,c}$ represents the spatial direction of $\text{UE}_{l,c}$.

4.2.2 Data transmission

The $\text{UE}_{l,c}$ receives the following signal

$$y_{l,c} = \sqrt{\eta} \mathbf{h}_{l,c} \mathbf{w}_c \mathbf{s}_c + \sum_{b \in \mathcal{C}, b \neq c} \sqrt{\eta} \mathbf{h}_{l,c} \mathbf{w}_b \mathbf{s}_b + n_{l,c}, \quad (4.2)$$

where $\mathbf{w}_c \in \mathbb{C}^{M \times 1}$ is the BF weight vector corresponding to the c -th cluster, $\mathbf{s}_c$ is the superposed signal of the c -th cluster and $n_{l,c} \sim \mathcal{N}(0, \sigma_n^2)$ is the additive white Gaussian noise at $\text{UE}_{l,c}$.

In (4.2), the first term represents the received superposed signal corresponding to the c -th cluster, while the second term is inter-cluster interference. Denote by $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_C] \in \mathbb{C}^{M \times C}$ the transmit beamforming matrix, then the normalization BF factor η is given by $\eta = \frac{1}{\text{Tr}(\mathbf{W}^H \mathbf{W})}$. To reduce the BF implementation complexity and overhead, we apply the AD-DBF technique studied in Chapter 3 that generates and steers a spatial beam toward each cluster, as shown in Fig. 4.1. Hence, $\mathbf{w}_c = \mathbf{a}_c$, where $\mathbf{a}_c = \mathbf{a}(\vec{\Theta}_c)$ is the array steering vector corresponding to the spatial direction $\vec{\Theta}_c$ of the c -th cluster, and thereby $\eta = \frac{1}{MC}$. The angle $\vec{\Theta}_c$ is calculated based on the spatial direction $\vec{\Theta}_{1,l,c}$ of UEs within the c -th cluster as follows

$$\vec{\Theta}_c = (\theta_c, \phi_c) = \begin{cases} \vec{\Theta}_{1,1,c} = (\theta_{1,1,c}, \phi_{1,1,c}) & \text{if } K_c = 1, \\ \left(\frac{\min_{l \in \mathcal{K}_c} \{\theta_{1,l,c}\} + \max_{l \in \mathcal{K}_c} \{\theta_{1,l,c}\}}{2}, \frac{\min_{l \in \mathcal{K}_c} \{\phi_{1,l,c}\} + \max_{l \in \mathcal{K}_c} \{\phi_{1,l,c}\}}{2} \right) & \text{if } K_c \geq 2. \end{cases} \quad (4.3)$$

The set $\Psi = \{\vec{\Theta}_{1,l,c}, l \in \mathcal{K}_c, c \in \mathcal{C}\}$ of the spatial directions of all UEs is called ADI. The main objective of this chapter is to design a MIMO-NOMA scheme with ADI feedback. This implies that the only channel information to be estimated is the set of spatial directions of UEs. This information can be estimated and tracked with standard AoDs estimation algorithms such as [119,120] at the BS or the receiver side. Throughout this dissertation, we assume a perfect estimation and feedback if needed. And, the impact of the imperfection of ADI on the proposed scheme is left for future work.

Using SC at the BS, NOMA superposes in power-domain the signals of UEs within the same cluster. The superposed signal $\mathbf{s}_c$ of the c -th cluster is thereby given by

$$\mathbf{s}_c = \sum_{l \in \mathcal{K}_c} \sqrt{\gamma_{l,c} p_c} s_{l,c}, \quad (4.4)$$

where $s_{l,c}$ is the modulated signal relative to $\text{UE}_{l,c}$, p_c is the power allocated to the c -th cluster, and $\gamma_{l,c}$ is the power allocation coefficient for $s_{l,c}$ such that $\sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1$.

4.2.3 AD MIMO-NOMA with SIC

Without loss of generality, let UEs in the c -th NOMA cluster be sorted in the descending order of their channel qualities, i.e., $\text{UE}_{l,c}$ is stronger than $\text{UE}_{t,c}$ ($\forall t > l \in \mathcal{K}_c$), where the user ordering strategy is detailed later in this subsection. Then, according to the NOMA principle, the power allocation coefficients assigned to UEs in the c -th cluster are sorted in the ascending order of their channel qualities $\gamma_{l,c}$ (to be defined later), i.e., $\gamma_{1,c} \leq \dots \leq \gamma_{K_c,c}$. Thus, using SIC decoding at the receiver side, every $\text{UE}_{l,c}$ can remove interference coming from the weaker UEs before decoding its own signal.

In the case of single-antenna NOMA, i.e., SISO-NOMA, the channel quality of $\text{UE}_{l,c}$ is the corresponding channel gain $|\mathbf{h}_{l,c}|$. Then, $|\mathbf{h}_{1,c}| \geq \dots \geq |\mathbf{h}_{K_c,c}|$. And, SIC works properly if the power allocation coefficients satisfy the following condition [121]

$$\frac{\gamma_{l,c} p_c |\mathbf{h}_{l-1,c}|^2}{\sigma_n^2} - \sum_{m=1}^{l-1} \frac{\gamma_{m,c} p_c |\mathbf{h}_{l-1,c}|^2}{\sigma_n^2} \geq P_{\min}, \quad l = 2, \dots, K_c, \quad (4.5)$$

where $P_{\min}$ is the minimum power difference between $\text{UE}_{l,c}$ and UEs with lower channel quality. Thus, SIC at $\text{UE}_{l,c}$ can decode first the signals of weak UEs with high allocated power, i.e., $\text{UE}_{t,c}$ ($\forall t > l$), and then subtract them from the superposed signal to decode its own signal.

In the case of multi-antenna NOMA with a single beam, using the ABF technique, the users are ordered according to their effective channel gain $|\mathbf{h}_{l,c} \mathbf{a}_c|^2$ or their distance w.r.t. the BS, both are considered in [86]. However, in the AD MIMO-NOMA system, UE also suffers from inter-cluster interference as seen in (4.2), unlike the single-antenna NOMA or the multi-antenna NOMA with a single beam. Therefore, the ordering strategy of the NOMA users and the constraint of efficient SIC in (4.5) must consider this interference. To this end, we generalize the channel quality $\zeta_{l,c}$ metric as the superposed-signal to other-clusters interference plus noise ratio at $\text{UE}_{l,c}$, as follows

$$\zeta_{l,c} = \frac{\eta p_c |\mathbf{h}_{l,c} \mathbf{a}_c|^2}{\sum_{b \in \mathcal{C}, b \neq c} \eta p_b |\mathbf{h}_{l,c} \mathbf{a}_b|^2 + \sigma_n^2}. \quad (4.6)$$

And, we order UEs such that $\zeta_{1,c} \geq \zeta_{2,c} \geq \dots \geq \zeta_{K_c,c}$. Besides, the constraint of efficient SIC

in the c -th cluster for MIMO-NOMA is given by

$$\gamma_{l,c}\zeta_{l-1,c} - \sum_{m=1}^{l-1} \gamma_{m,c}\zeta_{l-1,c} \geq P_{\min}, \quad l = 2, \dots, K_c. \quad (4.7)$$

This user ordering strategy thereby requires FCSI at the transmitter, which involves a high link overhead. For that, in Section 4.3, we propose another ordering strategy with only ADI feedback.

Assuming a perfect SIC decoding, the received signal $y_{l,c}$ at UE $_{l,c}$ in (4.2) can be rewritten as

$$y_{l,c} = \underbrace{\sqrt{\eta}\sqrt{\gamma_{l,c}p_c}\mathbf{h}_{l,c}\mathbf{a}_c s_{l,c}}_{\text{desired signal}} + \underbrace{\sum_{m=1}^{l-1} \sqrt{\eta}\sqrt{\gamma_{m,c}p_c}\mathbf{h}_{l,c}\mathbf{a}_c s_{m,c}}_{\text{intra-cluster interference}} + \underbrace{\sum_{b \in C, b \neq c} \sqrt{\eta}\sqrt{p_b}\mathbf{h}_{l,c}\mathbf{a}_b s_b}_{\text{inter-cluster interference}} + n_{l,c}. \quad (4.8)$$

Accordingly, UE $_{l,c}$ in the c -th NOMA cluster has the following SINR, while decoding its own message

$$\begin{aligned} \text{SINR}_{l,c}^{\text{NOMA}} &= \frac{\eta\gamma_{l,c}p_c|\mathbf{h}_{l,c}\mathbf{a}_c|^2}{\sum_{\substack{b \in C \\ b \neq c}} \eta p_b |\mathbf{h}_{l,c}\mathbf{a}_b|^2 + \sum_{m=1}^{l-1} \eta\gamma_{m,c}p_c |\mathbf{h}_{l,c}\mathbf{a}_c|^2 + \sigma_n^2}, \\ &= \frac{\gamma_{l,c}\zeta_{l,c}}{\sum_{m=1}^{l-1} \gamma_{m,c}\zeta_{l,c} + 1}. \end{aligned} \quad (4.9)$$

In (4.9), the first term in the denominator is inter-cluster interference, and the second one is the residual intra-cluster interference after SIC.

In the c -th SU cluster, one UE exists and has the following $\text{SINR}_{1,c}^{\text{SU}}$

$$\text{SINR}_{1,c}^{\text{SU}} = \frac{\eta p_c |\mathbf{h}_{1,c}\mathbf{a}_c|^2}{\sum_{b \in C, b \neq c} \eta p_b |\mathbf{h}_{1,c}\mathbf{a}_b|^2 + \sigma_n^2} = \zeta_{1,c}. \quad (4.10)$$

Considering a successful decoding at each UE and no propagation error within the total bandwidth B , UE $_{l,c}$ achieves the following throughput according to Shannon formula

$$R_{l,c} = B \log_2(1 + \text{SINR}_{l,c}). \quad (4.11)$$

Therefore, the system sum-throughput R can be expressed as

$$R = \sum_{c \in \mathcal{C}} \sum_{l \in \mathcal{K}_c} R_{l,c}, \quad (4.12a)$$

$$= \underbrace{\sum_{c \in \mathcal{C}_s} B \log_2(1 + \zeta_{1,c})}_{\text{SU clusters}} + \underbrace{\sum_{c \in \mathcal{C}_n} \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c} \zeta_{l,c}}{\sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l,c} + 1} \right)}_{\text{NOMA clusters}}. \quad (4.12b)$$

As seen in (4.12b), R depends on the PA coefficients $\boldsymbol{\gamma} = \{\gamma_{l,c}, c \in \mathcal{C}_n, l \in \mathcal{K}_c\}$ and on the UEs' channel qualities $\boldsymbol{\zeta} = \{\zeta_{l,c}, c \in \mathcal{C}, l \in \mathcal{K}_c\}$ built using FCSI $\mathbf{h}_{l,c}$ as shown in (4.6). Therefore, R can be improved by carefully designing a user clustering (UC) algorithm and a power allocation (PA) technique, including inter-cluster PA $\mathbf{p} = \{p_c, c \in \mathcal{C}\}$ and intra-cluster PA $\boldsymbol{\gamma}$ strategies. However, since we seek an AD MIMO-NOMA system only based on ADI, in the next section, we derive a channel quality metric (for user ordering strategy), a sum-throughput expression (for PA), and a spatial interference metric (for UC), all only based on $\boldsymbol{\Psi}$ the set of users' spatial directions.

4.3 ADI-based performance metrics

In this section, we derive ADI-based performance metrics: the user channel quality, the sum-throughput, and inter-user spatial interference, only based on ADI. These metrics will be used in the user ordering strategy in Section 4.3.1 and the design of UC and PA methods in Section 4.4 and 4.5, respectively.

4.3.1 ADI-based channel quality metric

The FCSI-based channel quality $\zeta_{l,c}$ in (4.6) is based on the channel vector $\mathbf{h}_{l,c}$, including both LoS and NLoS paths. However, AD-DBF doesn't exploit the NLoS paths and generates a single beam in the LoS path direction. To maintain low channel feedback, we consider a simplifying assumption that the NLoS path gains are negligible compared to those of the LoS path, i.e., $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \rightarrow 0, \forall 1 < q \leq N_{l,c}$. It is reasonable in mmWave channels since the transmission is highly directional at mmWave frequencies [115]. In addition, the authors in [122] show that the amplitudes of the NLoS paths are 20 dB weaker than the LoS path. Thereby, $\mathbf{h}_{l,c} \mathbf{a}_c$ can be simplified as

$$\mathbf{h}_{l,c} \mathbf{a}_c \xrightarrow{\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \rightarrow 0} \alpha_{1,l,c} \mathbf{a}_{1,l,c}^H \mathbf{a}_c. \quad (4.13)$$

(4.13) includes the complex gain $\alpha_{1,l,c}$ and the spatial direction $\vec{\Theta}_{1,l,c}$ of the LoS path. By substituting (4.13) in (4.6), we obtain the following channel quality $\check{\zeta}_{l,c}$ of UE $_{l,c}$

$$\zeta_{l,c} \xrightarrow{\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \rightarrow 0} \check{\zeta}_{l,c} = \frac{\eta p_c |\alpha_{1,l,c} \mathbf{a}_{1,l,c}^H \mathbf{a}_c|^2}{\sum_{b \in \mathcal{C}, b \neq c} \eta p_b |\alpha_{1,l,c} \mathbf{a}_{1,l,c}^H \mathbf{a}_b|^2 + \sigma_n^2}. \quad (4.14)$$

In high signal-to-noise ratio (SNR) regime, the channel quality $\check{\zeta}_{l,c}$ can be approximated by

$$\check{\zeta}_{l,c} \approx \frac{p_c |\mathbf{a}_{1,l,c}^H \mathbf{a}_c|^2}{\sum_{b \in \mathcal{C}, b \neq c} p_b |\mathbf{a}_{1,l,c}^H \mathbf{a}_b|^2}. \quad (4.15)$$

As seen in (4.15), $\check{\zeta}_{l,c}$ depends on inter-cluster PA strategy $\mathbf{p}$, the spatial direction of the considered UE $\vec{\Theta}_{1,l,c}$ and the angular information of each cluster $\vec{\Theta}_b$ ($\forall b \in \mathcal{C}$).

The prior works on MIMO-NOMA with limited feedback use the angular distance [123], the distance from the BS [86, 124], the large-scale fading coefficients [125] or the effective channel gains $|\mathbf{h}_{l,c} \mathbf{a}_c|^2$ [25, 99] as the user channel quality, to determine the order of SIC decoding. Different from these works, we leverage the highly directionality of the transmission properties and we propose a sub-optimal user ordering strategy that orders the users in the c -th NOMA cluster according to the following criterion

$$\check{\zeta}_{1,c} \geq \check{\zeta}_{2,c} \geq \dots \geq \check{\zeta}_{K_c,c}. \quad (4.16)$$

Interestingly, unlike prior works and our previous work [99], this strategy considers inter-cluster spatial interference, represented by the term in the denominator of (4.15). However, the distance or the path gain is not needed anymore. And, we will show in Section 4.6 that it is a winning strategy in the context of congested cells where the system is in an interference-limited regime.

4.3.2 ADI-based sum-throughput metric

The objective of this work is to design a PA method that maximizes the system sum-throughput with the ADI feedback. However, the objective function R in (4.12) requires FCSI, and hence, we obtain an FCSI-based PA. To develop an ADI-based PA, we replace (4.15) with (4.6) in (4.12) such that the ADI-based sum-throughput is defined as follows

$$\check{R} = \sum_{c \in \mathcal{C}_s} B \log_2(1 + \check{\zeta}_{1,c}) + \sum_{c \in \mathcal{C}_n} \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c} \check{\zeta}_{l,c}}{\sum_{m=1}^{l-1} \gamma_{m,c} \check{\zeta}_{l,c} + 1} \right). \quad (4.17)$$

We note that $\check{R}$ is obtained by considering the simplifying assumption $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \rightarrow 0$ in a high SNR regime. These assumptions will be verified by simulations in Section 4.6. $\check{R}$ depends on ADI Ψ and the PA technique, including inter-cluster PA $\mathbf{p}$ and intra-cluster PA γ strategies, which will

be described in Section 4.5.

4.3.3 ADI-based inter-user spatial interference metric

In this subsection, we focus on the system before clustering, and we use the subscript k ($\forall k \in \mathcal{K}$) as UE ID instead of (l, c) used after clustering. The objective is to define an ADI-based inter-user spatial interference metric needed for UC. Most of the existing works on AD MIMO-NOMA utilize the angular distance for UC. However, this metric does not describe inter-user interference properly when DBF is adopted. Indeed, the beamwidth is very dependent on the antennas' number and the steering angle of the beam, as explained in Section 2.2.3. Interestingly, we propose an inter-user spatial interference metric that includes the angular distance in addition to the main-beam width information.

For AD-DBF in a mono-path environment, the interference term at UE_k , observed at the denominator in (3.11), only depends on $\mathbf{a}_{1,k}^H \mathbf{a}_{1,u}$ ($\forall u \neq k \in \mathcal{K}$), i.e., the spatial interference causing by the beam generated for UE_u with the LoS path in the UE_k 's channel. Therefore, we define the normalized inter-user spatial interference metric $\beta_{k,u}$ in a mono-path environment as follows

$$\beta_{k,u} \stackrel{\text{def}}{=} \frac{1}{M} |\mathbf{a}_{1,k}^H \mathbf{a}_{1,u}|, \quad k, u \in \mathcal{K}, \quad (4.18)$$

where $\mathbf{a}_{1,u}$ is the array steering vector of the beam generated toward UE u spatial direction, i.e., $\vec{\Theta}_{1,u}$ and $\mathbf{a}_{1,k}^H$ is the array steering vector of the LoS path in UE k channel. Thus, $\beta_{k,u}$ is the normalized spatial interference of the beam generated toward UE u with the LoS path in UE k channel. By applying the simplifying assumption $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \rightarrow 0$ in the multi-path environment, we can use the same definition given by (4.18).

Using a $M_x \times M_z$ -URA array, the spatial interference $\beta_{k,u}$ in (4.18) can be rewritten based on (2.1) as follows

$$\beta_{k,u} = \frac{1}{M} \left| \sum_{m_z=1}^{M_z} \sum_{m_x=1}^{M_x} e^{j\{(m_x-1)(\omega_x(\vec{\Theta}_{1,u}) - \omega_x(\vec{\Theta}_{1,k})) + (m_z-1)(\omega_z(\vec{\Theta}_{1,u}) - \omega_z(\vec{\Theta}_{1,k}))\}} \right| \quad (4.19a)$$

$$= \left| \frac{\sin(M_z(\omega_z(\vec{\Theta}_{1,u}) - \omega_z(\vec{\Theta}_{1,k})))}{M_z \sin(\omega_z(\vec{\Theta}_{1,u}) - \omega_z(\vec{\Theta}_{1,k}))} \right| \left| \frac{\sin(M_x(\omega_x(\vec{\Theta}_{1,u}) - \omega_x(\vec{\Theta}_{1,k})))}{M_x \sin(\omega_x(\vec{\Theta}_{1,u}) - \omega_x(\vec{\Theta}_{1,k}))} \right| \quad (4.19b)$$

$\beta_{k,u}$ only depends on the spatial directions $\vec{\Theta}_{1,k}$ and $\vec{\Theta}_{1,u}$. According to (4.19b), $\beta_{k,u} = \beta_{u,k} \in (0, 1)$, and $\beta_{k,u}$ represents the value of the normalized array factor $AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta})$ of the beam pointed at the spatial direction $\vec{\Theta}_{1,u}$ for $\vec{\Theta} = \vec{\Theta}_{1,k}$ such that

$$\beta_{k,u} = |AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta}_{1,k})|. \quad (4.20)$$

Note that ULA along the x -axis is a particular case of URA with $M_x = M$, $M_z = 1$ and $\phi = 0$. Without loss of generality, with ULA, we simplify $AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta})$ to $AF_{(\theta_{1,u})}(\theta)$.

To illustrate (4.20), in Fig. 4.2, we depict the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ (blue curve)

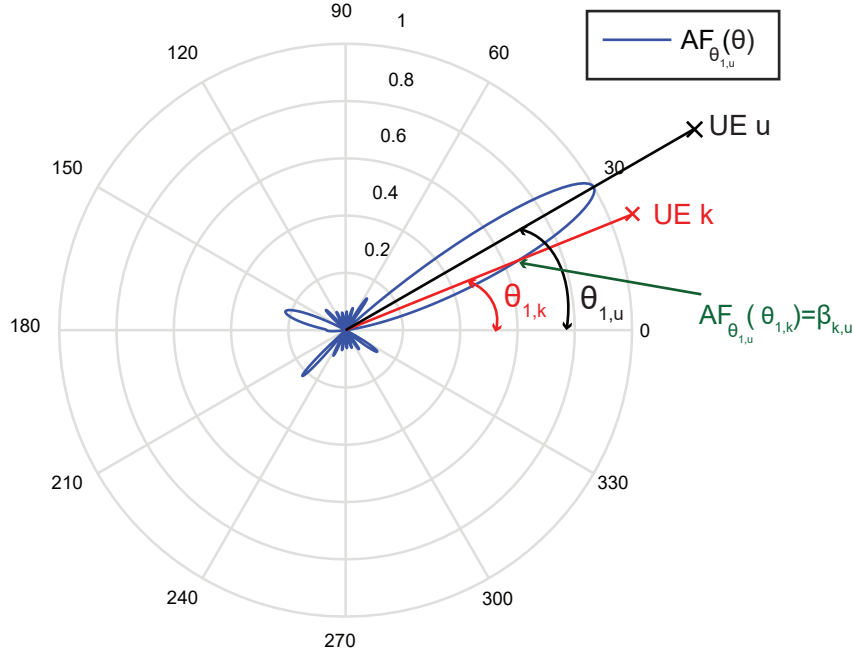


Figure 4.2: Illustration of the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ of the beam pointed toward UE u and the value $AF_{(\theta_{1,u})}(\theta_{1,k})$ for the spatial direction $\theta_{1,k}$ corresponding to $\beta_{k,u}$ of UE k and UE u .

of the beam pointed toward UE u and the corresponding value of $\beta_{k,u}$ when a horizontal ULA array is stacked along the x -axis. Hence, $\vec{\Theta} = (\theta, 0)$ and $M = M_x$.

Remark 1 Based on (4.19), we can see that if the angular distance $\Delta\vec{\Theta} = |\vec{\Theta}_{1,k} - \vec{\Theta}_{1,u}| \rightarrow (0, 0)$, then $\beta_{k,u} \rightarrow 1$. Thus, the smaller $\Delta\vec{\Theta}$ is, i.e., UE k and UE u are closely in the same direction, the larger $\beta_{k,u}$ is.

Remark 2 By referring to Subsection 2.2.3, $\Omega_u^{3dB} = |\vec{\Theta}_{1,u} - \vec{\Theta}|$ with $AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta}) = \sqrt{\frac{1}{2}}$ gives a measure of the 3dB-width of the beam generated at UE u . This 3dB-beamwidth Ω_u^{3dB} depends on the beam direction $\vec{\Theta}_{1,u}$ and the number of antennas M .

Thus, β is an important metric to determine the level of inter-user spatial interference using only the users spatial directions $\vec{\Theta}_{1,k}$ ($\forall k \in \mathcal{K}$), as depicted in Fig. 4.2. Moreover, we define the interference threshold β_0 , such that $\beta_{k,u} \geq \beta_0$ means that the LoS path of UE k lies in the UE u beam. Particularly, the β_0 -beamwidth $\Omega_u^{\beta_0}$ defines the angular distance $|\vec{\Theta}_{1,u} - \vec{\Theta}_0|$ satisfying $|AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta}_0)| = |AF_{(\vec{\Theta}_{1,u})}(\vec{\Theta}_{1,u} \pm \Omega_u^{\beta_0})| = \beta_0$. Using a fixed β_0 , $\Omega_u^{\beta_0}$ is varying w.r.t. the number of antennas and the spatial direction of UE u . Therefore, with the defined spatial interference metric β , we are able to define a unique threshold β_0 whatever the number of antennas or the users' spatial directions.

Lemma 1 UE k , having $\beta_{k,u} \geq \beta_0$, is located in the β_0 -width of the main lobe of the UE u beam, only if $\beta_0 > 0.217$.

Proof 1 With $d = \frac{\lambda}{2}$, a single main lobe exists in $(0, \pi)$, surrounded by a number of side lobes [30]. $\beta_{k,u} \geq \beta_0$ ensures that UE k belongs to $\Omega_u^{\beta_0}$. However, $\Omega_u^{\beta_0}$ can be associated to the

main lobe or one of the side lobes. To select only the users covered by the main lobe, β_0 must be lower bounded by the first-side lobe level β^{FSL} , i.e., $\beta_0 > \beta^{FSL}$. When the BS is equipped with a uniformly excited phased array, this value is given by $\beta^{FSL} = 0.217$ or (-13.26 dB) for any values of the antennas number M and the steering angle $\theta_{1,u}$ [30], when $d = \frac{\lambda}{2}$, as illustrated by Fig. 4.3. Here, Lemma 1 is proved.

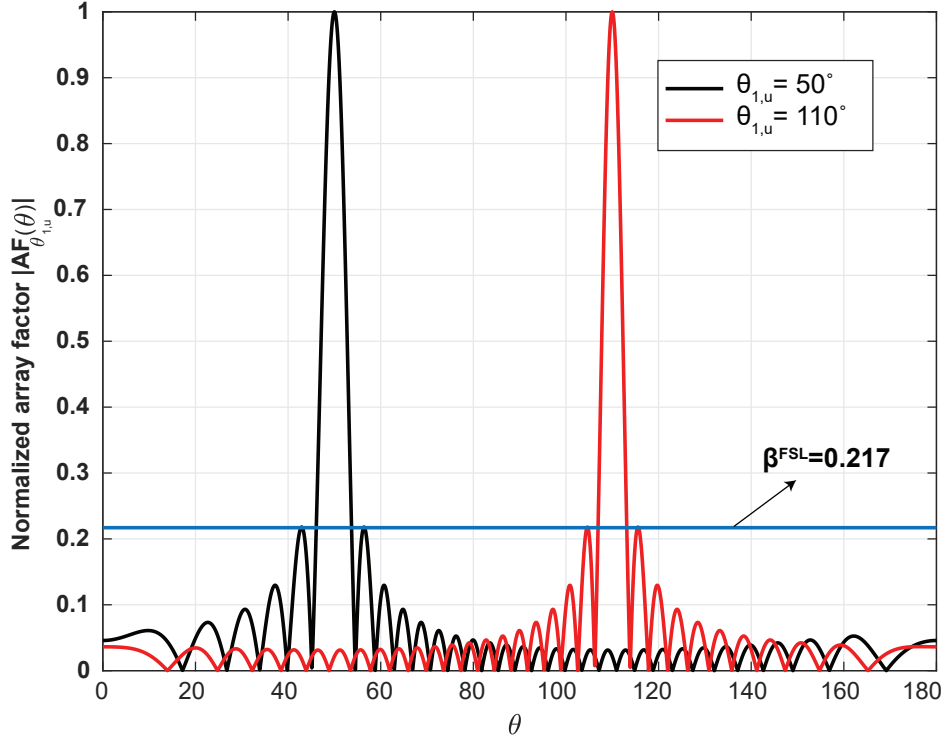


Figure 4.3: Illustration of the normalized array factor $|AF^n_{(\theta_{1,u})}(\theta)|$ of the beam generated toward UE u for $\theta_{1,u} \in \{50^\circ, 110^\circ\}$, with $M = 32$.

We will use this metric in the design of the UC algorithm described in Section 4.4. The value 0.217 in Lemma 1 will be observed by simulations in Section 4.6.

4.4 User clustering algorithm

The UC optimization problem that maximizes the sum-throughput R in (4.12) is a non-convex problem, which can only be solved using an exhaustive search. However, that would result in an exponential complexity, brought by calculating the sum-throughput among all the possible sets of clusters. Indeed, the NOMA cluster size is not restricted in the case of multi-user AD MIMO-NOMA, i.e., $1 \leq K_c \leq K$, hence the number of all the possible combinations increases exponentially with the number of users, K . The previous UC methods using FCSI or partial CSI [126–129], such as the large-scale fading coefficients or the channel correlation, are based on the K-means or agglomerative algorithms, which require the target number of clusters as an input parameter. However, for practical applications, this parameter is a priori unknown at the BS and an arbitrary value may lead to suboptimal solutions. To reduce the exponential complexity

of the exhaustive search method on the one hand, and on the other hand the CSI overhead and the number of input parameters, we leverage the potentiality of the ADI in the highly directional mmWave transmission. Specifically, we exploit the defined ADI-based inter-user spatial interference metric to group the users within SU and NOMA clusters. And, the proposed UC algorithms aim to reduce inter-user interference and thus improve the sum throughput. To the best of our knowledge, this metric is used for the first time in the context of UC. Next, we design a low-complex 2-user clustering algorithm that groups the users between SU and 2-user NOMA clusters. Furthermore, we develop a general clustering algorithm to allow for multi-user communication per NOMA cluster.

Let $\mathcal{G}_c$ be the set of UE IDs in $\mathcal{K}$ belonging to the c -th cluster. Accordingly, if $\mathcal{G}_c\{l\} = k$, then UE $k \in \mathcal{K}$ is the l -th UE in the c -th cluster. We assume that each user belongs to only one cluster, thereby $\mathcal{G}_c \cap \mathcal{G}_{c'} = \emptyset, \forall c \neq c' \in \mathcal{C}$ and $\cup_{c=1}^C \mathcal{G}_c = \mathcal{K}$. Let β_0 be the inter-user spatial interference threshold, as defined in Section 4.3.3. We are interested in the two UEs having $\beta_{k,u} \geq \beta_0$. Hence, we first select the set of groups of 2-users having a spatial interference greater than β_0 in the following set $\mathcal{D}$

$$\mathcal{D} = \{(k, u), \beta_{k,u} \geq \beta_0, k < u \in \mathcal{K}\}. \quad (4.21)$$

4.4.1 2-user clustering algorithm

Given the set ADI Ψ , the main idea of the proposed 2-user UC, denoted as 2-user β -UC, is to group two-by-two the UEs (e.g., UE_u and UE_k) with a significant inter-user interference, i.e., $\beta_{k,u} > \beta_0$, into the same NOMA cluster, whose beam angle is defined in (4.3). Subsequently, we assign a single beam to each one of the remaining UEs with a cluster angle equal to their LoS path angle, as mentioned in (4.3).

The starting point for the first 2-user NOMA cluster is the point (o, r) with the largest value in $\mathcal{D}$, i.e., $\beta_{o,r} \geq \beta_0$. Then, we assign one beam ($c = 1$) to UE_o and UE_r , with $\vec{\Theta}_c = \frac{\vec{\Theta}_{1,o} + \vec{\Theta}_{1,r}}{2}$. Subsequently, we remove these two UEs from the set $\mathcal{D}$ to prevent their existence in another cluster. We then repeat this stage until $\mathcal{D}$ is empty. Finally, we assign a single cluster to each of the remaining UEs with a cluster angle equal to the angle of their LoS path, as noted in (4.3). The proposed 2-user β -UC algorithm is summarized in Algorithm 1.

4.4.2 Multi-user clustering algorithm

The proposed multi-user UC algorithm, denoted as multi-UE β -UC, aims to find the set $\mathcal{G}_c \forall c \in \mathcal{C}$, so that all UEs having interference with each others are grouped in the same cluster. To do so, we first find the NOMA clusters by using (4.22), while starting by the pair of UEs (o, r) having the largest spatial interference such that $\beta_{o,r} \geq \beta_0$. Specifically, this phase is a nested loop, where the outer loop finds the C_n NOMA clusters, and the inner loop searches the users belonging to the current c -th NOMA cluster. Once all the NOMA clusters are obtained, we assign to each one of the remaining UEs a single cluster as mentioned in (4.23). The beam angle associated to each cluster is calculated as in (4.3). More details of the proposed β -UC algorithm

where $\mathcal{F}_{p,c}$ denotes the set $\mathcal{F}$ corresponding to the c -th and p -th iterations of the outer and the inner loop, respectively. Note that there is no inner loop ($\lfloor \frac{K_c}{2} \rfloor = 1$) in the 2-user β -UC algorithm.

Lemma 2 *Once β -UC is applied with $\beta_0 > \beta^{FSL}$, UEs in the c -th NOMA cluster are located within the β_0 -width of the main lobe corresponding to the c -th NOMA cluster, i.e., $\beta_{g,c} \stackrel{def}{=} |AF_{(\vec{\Theta}_c)}(\vec{\Theta}_{1,g})| \geq \beta_0 \forall g \in \mathcal{G}_c$.*

Proof 2 *Without loss of generality, we prove Lemma 2 using ULA with $\vec{\Theta} = \theta$. And this proof can be easily extended for URA with $\vec{\Theta} = (\theta, \phi)$.*

The set $\mathcal{G}_c$ of UE IDs within the c -th NOMA cluster satisfies (4.22), thereby $\beta_{g,d} \geq \beta_0 \forall d \neq g \in \mathcal{G}_c$. Let $\min_{g \in \mathcal{G}_c} \{\theta_{1,g}\} = \theta_{1,v}$ and $\max_{g \in \mathcal{G}_c} \{\theta_{1,g}\} = \theta_{1,w}$, then the spatial angle θ_c corresponding to the c -th NOMA cluster is equal to $\theta_c = \frac{\theta_{1,v} + \theta_{1,w}}{2}$, as defined in (4.3).

We define $\beta_{v,c} \stackrel{def}{=} |AF_{(\theta_c)}(\theta_{1,v})|$, which can be rewritten by $\beta_{v,c} = |AF_{(\theta_{1,v})}(\theta_c)|$, with $|AF_{(\theta_{1,v})}(\theta)|$ is the normalized array factor of the beam steered toward UE v . Now, we aim to demonstrate that $|AF_{(\theta_{1,v})}(\theta_c)| \geq \beta_0$. Since $\beta_{v,w} \geq \beta_0$ and $\beta_0 > \beta^{FSL}$, then UE w is covered by the β_0 -width of the main lobe of the UE v beam according to Lemma 1. Therefore, and by exploiting the property of the function in (4.19b), $|AF_{(\theta_{1,v})}(\theta)|$ is a decreasing function of θ in the angle interval $I = [\theta_{1,v}, \theta_{1,w}]$. Accordingly, $|AF_{(\theta_{1,v})}(\theta_c)| > |AF_{(\theta_{1,v})}(\theta_{1,w})| \geq \beta_0$.

Similarly, we can demonstrate that $|AF_{(\theta_{1,w})}(\theta_c)| \geq \beta_0$.

However, $|AF_{(\theta_c)}(\theta_{1,v})| = |AF_{(\theta_{1,v})}(\theta_c)| \geq \beta_0$ and $|AF_{(\theta_c)}(\theta_{1,w})| = |AF_{(\theta_{1,w})}(\theta_c)| \geq \beta_0$. Generally, $\theta_{1,v} \leq \theta_{1,g} \leq \theta_{1,w} \forall g \in \mathcal{G}_c$, and $|AF_{(\theta_c)}^n(\theta)|$ has one maximum at $\theta = \theta_c$ in the interval I . This implies that $|AF_{(\theta_c)}(\theta_{1,g})| \geq \min\{|AF_{(\theta_c)}(\theta_{1,v})|, |AF_{(\theta_c)}(\theta_{1,w})|\} \geq \beta_0 \forall g \in \mathcal{G}_c$. In other words, UE g is located within the β_0 -width of the main beam corresponding to the c -th NOMA cluster, according to Lemma 1. Here, Lemma 2 is proved.

If the condition in Lemma 1 is met, i.e., $\beta_{k,u} \geq \beta_0$ with $\beta_0 > \beta^{FSL} = 0.217$, Lemma 2 ensures that the choice of the cluster angle in (4.3) guarantees that the users are covered by the main lobe corresponding to the c -th NOMA cluster within the β_0 -width. This in turn ensures that the beam misalignment is not severe.

4.5 Multi-user power allocation technique

In this section, we consider FCSI- and ADI-based PA problems that maximize the FCSI- and the ADI-based sum-throughput, respectively. Throughout this section, we propose a fixed inter-cluster PA strategy to obtain a convex optimization problem. Then, the closed-form intra-cluster PA can be given using the KKT conditions. To the best of our knowledge, this is the first work that utilizes the ADI-based sum-throughput to design a PA policy with ADI feedback.

4.5.1 Problem formulation

Given the set $\mathcal{G}_c$ ($\forall c \in \mathcal{C}$), we aim to design a PA technique that maximizes the system sum-throughput of the multi-user AD MIMO-NOMA scheme. The PA technique consists to allocate

Algorithm 2: β -based User Clustering Algorithm**Input:** $\theta_{1,k}, \beta_{k,u} \forall k < u \in \mathcal{K}, \beta_0$.**Output:** $C_n, C, \mathcal{G}_c \forall c \in C$.*Initialization* : $c = 0$ 1: Select the set of groups of 2-users having a spatial interference greater than β_0 s.t. $\mathcal{D} = \{(k, u), \beta_{k,u} \geq \beta_0, k < u \in \mathcal{K}\}$.*Phase 1: Clustering of users with high inter-user spatial interference in NOMA clusters*2: **repeat**3: $c = c + 1, \mathcal{F} = \mathcal{D}, \mathcal{G}_c = \{\}$;4: **repeat**5: Locate in $\mathcal{F}$ the group of 2-users having the largest spatial interference: $\beta_{o,r} = \max_{(k,u) \in \mathcal{F}} \{\beta_{k,u}\}, \mathcal{G}_c = \{\mathcal{G}_c, o, r\}$.6: Remove from $\mathcal{D}$ the couple (o, r) and their coupling with any other UEs to prevent their presence in another cluster: $\mathcal{D} \leftarrow \mathcal{D} - \{(o, r)\} - \{(o, w)\} - \{(v, o)\} - \{(r, m)\} - \{(t, r)\}$
($\forall w > o, v < o, m > r, t < r \in \mathcal{K}$)7: Select the set of UEs in interference with UE o : $\mathcal{V} = \{v : (o, v) \in \mathcal{F}, o < v \text{ or } (v, o) \in \mathcal{F}, o > v\}$.8: Select the set of UEs in interference with UE r : $\mathcal{W} = \{w : (r, w) \in \mathcal{F}, r < w \text{ or } (w, r) \in \mathcal{F}, r > w\}$ 9: Select UEs in interference with both UE $_o$ and UE $_r$: $\mathcal{Y} = \mathcal{V} \cap \mathcal{W}$ 10: **if** $|\mathcal{Y}| == 1$ **then**11: // One user exists in $\mathcal{F}$ Update $\mathcal{F}$ and $\mathcal{G}_c$: $\mathcal{F} = \emptyset, y = \mathcal{Y}, \mathcal{G}_c = [\mathcal{G}_c, y],$ 12: Remove from $\mathcal{D}$ the coupling of UE y with any other UEs to prevent their presence in another cluster: $\mathcal{D} \leftarrow \mathcal{D} - \{(y, w)\} - \{(v, y)\} (\forall w > y, v < y \in \mathcal{K})$.13: **else**14: // Multiple users exist in $\mathcal{F}$ Update $\mathcal{F}$ and select the set of groups of 2-users in $\mathcal{Y}$ who have a spatial interference greater than β_0 : $\mathcal{F} = \{(k, u), \beta_{k,u} \geq \beta_0, k < u \in \mathcal{Y}\} - \{(o, r)\}$.15: **end if**16: **until** $\mathcal{F} = \emptyset$ 17: **until** $\mathcal{D} = \emptyset$ $C_n = c$.*Phase 2: Clustering of the remaining users in SU clusters*18: **for** $k = 1 : K$ **do**19: **if** $k \notin \cup_{l=1}^c \mathcal{G}_l$ **then**20: $c = c + 1; \mathcal{G}_c = \{k\}$.21: **end if**22: **end for** $C = c$.

the power between the clusters $\mathbf{p}$ and between UEs in each NOMA cluster $\boldsymbol{\gamma}$. By defining P_c as the emitted power toward the c -th cluster as follows

$$P_c = \eta \|\mathbf{w}_c\|^2 \sum_{l=1}^{K_c} \gamma_{l,1} p_c = \eta \|\mathbf{w}_c\|^2 p_c, \quad (4.25)$$

the optimization problem can be formulated as follows

$$(\mathcal{P}_1) : \quad \{\boldsymbol{\gamma}^*, \mathbf{p}^*\}^{\text{type}} = \max_{\boldsymbol{\gamma}, \mathbf{p}} R^{\text{type}}(\boldsymbol{\gamma}, \mathbf{p}), \quad (4.26)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1, \quad \forall c \in \mathcal{C}, \quad (4.26a)$$

$$\sum_{c \in \mathcal{C}} P_c \leq P_e, \quad (4.26b)$$

$$\delta(\boldsymbol{\gamma}, \zeta_{l-1,c}^{\text{type}}) \geq P_{\min}, \quad 2 \leq l \leq K_c, \quad \forall c \in \mathcal{C}_n. \quad (4.26c)$$

where $\text{type} = \{\text{FCSI}, \text{ADI}\}$, $R^{\text{FCSI}} = R$ in (4.12), $R^{\text{ADI}} = \check{R}$ in (4.17), $\zeta_{l,c}^{\text{FCSI}} = \zeta_{l,c}$ in (4.6), $\zeta_{l,c}^{\text{ADI}} = \check{\zeta}_{l,c}$ in (4.15), P_e is the total transmit power and $\delta(\boldsymbol{\gamma}, \zeta_{l-1,c}^{\text{type}}) = \gamma_{l,c} \zeta_{l-1,c}^{\text{type}} - \sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l-1,c}^{\text{type}}$. (4.26a) and (4.26b) are respectively the total intra-cluster and inter-cluster PA constraints, while (4.26c) is the SIC constraint given by (4.7). $(\mathcal{P}_1)$ maximizes the system sum-throughput by jointly optimizing intra- and inter-cluster PA. However, the PA factors from different UEs are very coupled in (4.26). To obtain closed-form solutions, we apply a fixed inter-cluster PA, i.e., the c -th cluster is assigned with a predefined fixed power p_c , as detailed in Section 4.5.2. Hence, $(\mathcal{P}_1)$ is reformulated as an intra-cluster PA maximization problem subject to the intra-cluster PA and efficient SIC constraints as follows

$$(\mathcal{P}_2) : \quad \{\boldsymbol{\gamma}^*\}^{\text{type}} = \max_{\boldsymbol{\gamma}} R^{\text{type}}(\boldsymbol{\gamma}), \quad (4.27)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1, \quad \forall c \in \mathcal{C} \quad (4.27a)$$

$$\delta(\boldsymbol{\gamma}, \zeta_{l-1,c}^{\text{type}}) \geq P_{\min}, \quad 2 \leq l \leq K_c, \quad \forall c \in \mathcal{C}_n. \quad (4.27b)$$

Denote R_c^{type} as the sum-throughput of UEs within the c -th cluster. For SU cluster, $R_c^{\text{type}} = B \log_2(1 + \zeta_{1,c}^{\text{type}}) = \text{constant} \quad \forall \boldsymbol{\gamma}$. For NOMA cluster, $R_c^{\text{type}} = \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c} \zeta_{l,c}^{\text{type}}}{\sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l,c}^{\text{type}} + 1} \right)$ depends only on the PA coefficients $\boldsymbol{\gamma}_c = \{\gamma_{l,c}, l \in \mathcal{K}_c\}$ of UEs in the c -th cluster. And, $\delta(\boldsymbol{\gamma}_c, \zeta_{l-1,c}^{\text{type}}) = \delta(\boldsymbol{\gamma}, \zeta_{l-1,c}^{\text{type}})$. Consequently, solving $(\mathcal{P}_2)$ is equivalent to solve $\mathcal{C}_n$ problems, independently. Therefore, for each NOMA cluster, the sum-throughput optimization problem

defining the intra-cluster PA strategy $\{\gamma_c^\star\}^{\text{type}} \forall c \in \mathcal{C}_n$ can be modeled as

$$(\mathcal{P}_3) : \{\gamma_c^\star\}^{\text{type}} = \max_{\gamma_c} \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c} \zeta_{l,c}^{\text{type}}}{\sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l,c}^{\text{type}} + 1} \right), \quad (4.28)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1, \quad (4.28a)$$

$$\delta(\gamma_c, \zeta_{l-1,c}^{\text{type}}) \geq P_{\min}, \quad 2 \leq l \leq K_c. \quad (4.28b)$$

After designing a fixed inter-cluster PA strategy in Section 4.5.2, we aim in Section 4.5.3 to design an intra-cluster PA strategy by solving $(\mathcal{P}_3)$ using the KKT conditions.

4.5.2 Fixed inter-cluster power allocation

With a uniform PA per cluster, i.e., $p_c = \text{constant} (\forall c)$, the inter-cluster power was allocated such that $\sum_{c=1}^C P_c = P_e$. This implied that the uniform PA per cluster satisfies $P_c = \frac{P_e}{C}$, using (4.25) and $\eta \|\mathbf{w}_c\|^2 = \frac{1}{C}$. In multi-user AD MIMO-NOMA scheme, the number of UEs per cluster is variable. So, we allocate a power p_c to the c -th cluster proportional to the number K_c of UEs served in the c -th cluster. Accordingly, based on (4.25), P_c is also proportional to K_c . This is supposed to guarantee a certain power fairness among UEs. Therefore, we calculate P_c as follows

$$P_c = K_c \frac{P_e}{K}. \quad (4.29)$$

From (4.25) and (4.29) we obtain that

$$p_c = K_c C \frac{P_e}{K}. \quad (4.30)$$

Thus, the power is non uniformly allocated to the clusters.

4.5.3 FCSI- and ADI-based intra-cluster power allocation

We apply a similar methodology as in [121]² to find the closed-form solution of $(\mathcal{P}_3)$, since it is convex under the constraints (4.28a) and (4.28b). Thus, the corresponding Lagrange function

²In [121], the authors consider a single-antenna NOMA system in both uplink and downlink. Given the set of NOMA clusters, they derive a PA policy that maximizes the sum-throughput per NOMA cluster under three constraints: (1) minimum data rate requirements, (2) transmission power and (3) efficient SIC constraints. The closed-form solutions for optimal PA are derived using KKT conditions for any cluster size.

$\mathcal{L}^{\text{type}}$ is defined as follows

$$\mathcal{L}^{\text{type}}(\gamma_c, \rho, \chi) = B \sum_{l=1}^{K_c} \log_2 \left(1 + \frac{\gamma_{l,c} \zeta_{l,c}^{\text{type}}}{\sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l,c}^{\text{type}} + 1} \right) + \rho \left(1 - \sum_{l=1}^{K_c} \gamma_{l,c} \right) + \sum_{l=2}^{K_c} \chi_l \left\{ \gamma_{l,c} \zeta_{l-1,c}^{\text{type}} - \sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l-1,c}^{\text{type}} - P_{\min} \right\} \quad (4.31)$$

where ρ and $\chi = \{\chi_l, l = 2, \dots, K_c\}$ are the Lagrange multipliers corresponding to the constraints (4.28a) and (4.28b), respectively, such that $\rho > 0$ and $\chi_l \geq 0$ ($l = 2, \dots, K_c$). The optimal PA solution satisfies the following KKT conditions

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \gamma_{1,c}} = \frac{B \zeta_{1,c}^{\text{type}}}{\gamma_{1,c} \zeta_{1,c}^{\text{type}} + 1} - \sum_{m=2}^{K_c} \frac{B \gamma_{m,c} (\zeta_{m,c}^{\text{type}})^2}{(\sum_{q=1}^m \gamma_{q,c} \zeta_{m,c}^{\text{type}} + 1) (\sum_{q'=1}^{m-1} \gamma_{q',c} \zeta_{m,c}^{\text{type}} + 1)} - \rho - \sum_{m=2}^{K_c} \chi_m \zeta_{m-1,c}^{\text{type}} = 0, \text{ if } \gamma_{1,c} \geq 0, \quad (4.32)$$

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \gamma_{l,c}} = \frac{B \zeta_{l,c}^{\text{type}}}{\sum_{m=1}^l \gamma_{m,c} \zeta_{l,c}^{\text{type}} + 1} - \rho + \chi_l (\zeta_{l-1,c}^{\text{type}}) - \sum_{m=l+1}^{K_c} \frac{B \gamma_{m,c} (\zeta_{m,c}^{\text{type}})^2}{(\sum_{q=1}^m \gamma_{q,c} \zeta_{m,c}^{\text{type}} + 1) (\sum_{q'=1}^{m-1} \gamma_{q',c} \zeta_{m,c}^{\text{type}} + 1)} - \sum_{m=l+1}^{K_c} \chi_m (\zeta_{m-1,c}^{\text{type}}) = 0, \text{ if } \gamma_{l,c} \geq 0, \forall l = 2, \dots, K_c, \quad (4.33)$$

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \rho} = f(\gamma) = 1 - \sum_{l=1}^{K_c} \gamma_{l,c} = 0, \text{ if } \rho > 0, \quad (4.34)$$

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \chi_l} = g_l(\gamma) = \gamma_{l,c} \zeta_{l-1,c}^{\text{type}} - \sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l-1,c}^{\text{type}} - P_{\min} \geq 0, \text{ if } \chi_l \geq 0, \forall l = 2, \dots, K_c, \quad (4.35)$$

$$\chi_l g_l(\gamma) = 0, \forall l = 2, \dots, K_c. \quad (4.36)$$

As seen in $(\mathcal{P}_3)$, we have $K_c - 1$ inequality constraints, leading to 2^{K_c-1} combinations of Lagrange multipliers as stated in (4.36). Note that $\gamma_{l,c} > 0 \forall l \in \mathcal{K}_c$ in our problem, then obtaining optimal feasible solutions for K_c decision variables requires exactly K_c equations [130]. This implies that only one combination can be checked to find a solution for the problem $(\mathcal{P}_3)$, i.e., $\chi_l > 0$ and $g_l(\gamma) = 0 \forall l \in \{2, \dots, K_c\}$. This is one of the several combinations in [121], where only the equations of the total power and SIC constraints are considered to find the optimal solutions. In other words, $f(\gamma) = 0$ and $g_l(\gamma) = 0 (\forall l)$, provide the optimal solution, while the other constraints provide the necessary conditions for the obtained optimal solution. Thereby, the

closed-form optimal solutions can be calculated similarly to [121] by

$$\gamma_{l,c}^{\star \text{type}} = \begin{cases} \frac{1}{2^{K_c-1}} - \sum_{m=1}^{K_c-1} \frac{P_{\min}}{2^m \zeta_{m,c}^{\text{type}}}, & \text{if } l = 1, \\ \frac{1}{2^{K_c-l+1}} + \frac{P_{\min}}{\zeta_{l-1,c}^{\text{type}}} - \sum_{m=l}^{K_c} \frac{P_{\min}}{2^{m-l+1} \zeta_{m-1,c}^{\text{type}}}, & \text{if } l \neq 1 \in \mathcal{K}_c. \end{cases} \quad (4.37)$$

For the dynamic intra-cluster PA solutions in [121], the authors verify that the corresponding set of Lagrange multipliers satisfies the KKT conditions for the different combinations. Therefore, since (4.37) is one of those in [121], the corresponding set of Lagrange multipliers $\{\rho, \chi\}$ also satisfies the KKT conditions. Unlike our previous work [99] applying a fixed intra-cluster PA for the two NOMA users, the proposed PA policy dynamically maximizes the sum-throughput per each NOMA cluster under the SIC constraints. The closed-form solution in (4.37) for the problem $(\mathcal{P}_3)$ is obtained with a linear complexity w.r.t. the number of all NOMA users, i.e., $O(\sum_{c=1}^{C_n} K_c)$.

4.5.4 Performance comparison: FCSI- versus ADI-based schemes

In this subsection, we compare the realistically achievable sum-throughput R_{Σ}^{FCSI} and R_{Σ}^{ADI} of the multi-user AD MIMO-NOMA schemes using FCSI or ADI feedbacks, respectively. They are expressed as follow

$$\begin{cases} R_{\Sigma}^{\text{FCSI}} = R(\{\gamma^{\star}\}^{\text{FCSI}}), \\ R_{\Sigma}^{\text{ADI}} = R(\{\gamma^{\star}\}^{\text{ADI}}), \end{cases} \quad (4.38)$$

$$(4.39)$$

where $R(\gamma)$ is the realistic sum-throughput given by (4.12), $\{\gamma^{\star}\}^{\text{FCSI}}$ and $\{\gamma^{\star}\}^{\text{ADI}}$ are calculated using (4.37).

UC and inter-cluster PA are identical for both sum-throughput. The difference is in the user ordering strategy and the computation of intra-cluster PA within each NOMA cluster. Likewise, the FCSI- and ADI-based sum-throughput of UEs within the c -th cluster are given by $R_c^{\text{FCSI}} = R_c(\{\gamma_c^{\star}\}^{\text{FCSI}})$ and $R_c^{\text{ADI}} = R_c(\{\gamma_c^{\star}\}^{\text{ADI}})$, respectively, with $R_c(\gamma_c) = \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c} \zeta_{l,c}^{\text{FCSI}}}{\sum_{m=1}^{l-1} \gamma_{m,c} \zeta_{l,c}^{\text{FCSI}+1}} \right)$. Obviously, $R_c(\{\gamma_c^{\star}\}^{\text{FCSI}}) \geq R_c(\{\gamma_c^{\star}\}^{\text{ADI}}) \forall c \in C_n$. The sum-throughput gain ΔR_{Σ} defined as $\Delta R_{\Sigma} = R_{\Sigma}^{\text{FCSI}} - R_{\Sigma}^{\text{ADI}}$ can be written as

$$\Delta R_{\Sigma} = \sum_{c \in C_n} \left(R_c(\{\gamma_c^{\star}\}^{\text{FCSI}}) - R_c(\{\gamma_c^{\star}\}^{\text{ADI}}) \right) \geq 0. \quad (4.40)$$

The gain ΔR_{Σ} only results from the sum-throughput change within the NOMA clusters due to different intra-cluster PA and user ordering strategies. Consequently, ΔR_{Σ} rises as the number of NOMA clusters grows.

In this work, we apply the proposed schemes in a mmWave rural environment, using NYUSIM. Thus, a single LoS path exists solely or with another NLoS path [22]. Thereby, we analyze ΔR_{Σ}

for the two following cases: (i) mono-path environment, i.e., $N_{l,c} = 1$ and (ii) multi-path environment with $N_{l,c} = 2$.

4.5.4.1 Mono-path environment

In mono-path environments, $\mathbf{h}_{l,c}\mathbf{a}_c$ is given by (4.13). And, in a high SNR regime, we obtain that $\zeta^{\text{ADI}} \approx \zeta^{\text{FCSI}}$. Thereby, the same user ordering strategy is applied using FCSI or ADI feedbacks, and we obtain $\star^{\text{ADI}} \approx \star^{\text{FCSI}}$ according to (4.37) and $\Delta R_\Sigma \approx 0$. Consequently, the proposed schemes have similar performance in a mono-path environment, using the channel vector $\mathbf{h}_{l,c}$ information or only ADI Ψ .

4.5.4.2 Multi-path environment

In multi-path environments, we applied the simplifying assumption $\frac{\alpha_{2,l,c}}{\alpha_{1,l,c}} \rightarrow 0$. In practice, $\frac{\alpha_{2,l,c}}{\alpha_{1,l,c}} = \epsilon_0$, with $0 < |\epsilon_0| < 1$, which in turn implies that $\zeta^{\text{ADI}} \neq \zeta^{\text{FCSI}}$, and $\star^{\text{ADI}} \neq \star^{\text{FCSI}}$. In a high SNR regime, ΔR only depends on ϵ_0 . If $\epsilon_0 \rightarrow 0$, then $\Delta R_\Sigma \rightarrow 0$. Therefore, in multi-path environments, the performance of ADI-based AD MIMO-NOMA scheme w.r.t. FCSI-based AD MIMO-NOMA depends on the complex gain of the NLoS path, if it exists, against that of the LoS path.

4.6 Illustrative results and discussions

FCSI and ADI are assumed to be perfectly known at the transmitter. We consider a rural multi-path channel model generated by NYUSIM. The channel and system parameters are listed in Table 4.1. The noise power in Table 4.1 represents the thermal noise power at a temperature of 25°C. We assume that the users belong to the same side of the BS, and their angles range from 10° to 165° to prevent the endfire beamforming.

Table 4.1: Simulation parameters

Carrier frequency	28 GHz
Channel bandwidth, B	20 MHz
Cell edge radius	100 m
Transmission power, P_e	30 dBm
Noise power, σ_n^2	-101 dBm
Minimum power difference, $P_{\min}$	1 mW
Number of paths per time cluster in rural environment	{1, 2}

4.6.1 Channel estimation overhead

To assess the channel estimation overhead of the simulated schemes, we provide the number of real coefficients needed per channel estimation for each user in Table 4.2. It is clear that ADI-based schemes require $2M$ fewer coefficients than those based on FCSI since only the user spatial direction is necessary.

	Number of real coefficients per channel estimation for each user
C-AD-DBF	N_e
AD MIMO-OMA	N_e
2-user AD MIMO-NOMA with FCSI NOMA	$2 \times M + N_e$ [$\mathbf{h}_k \in \mathbb{C}^{1 \times M}, \theta_{1,k}$]
2-user AD MIMO-NOMA with ADI NOMA	N_e
multi-user AD MIMO-NOMA with FCSI NOMA	$2 \times M + N_e$ [$\mathbf{h}_k \in \mathbb{C}^{1 \times M}, \theta_{1,k}$]
multi-user AD MIMO-NOMA with ADI NOMA	N_e

Table 4.2: Channel Estimation Overhead. Note that $N_e = 1[\theta_{1,k}]$, (resp. $2[(\theta_{1,k}, \phi_{1,k})]$) with ULA (resp. URA).

4.6.2 2D AD MIMO-NOMA performance

In this section, we investigate the performance evaluation of the different proposed techniques when the BS adopts a ULA array with $M = 32$ antennas stacked horizontally along the x -axis to serve K users in a rural multi-path environment.

4.6.2.1 Performance of user ordering strategy

In Fig. 4.4, the impact of various user ordering strategies using FCSI or limited feedback on the 2-user AD MIMO-NOMA spectral efficiency (sum-throughput per bandwidth) is presented. The optimal strategy orders the users w.r.t. their realistic channel qualities ζ , which is given in (4.6), such that $\zeta_{1,c} \geq \dots \geq \zeta_{K,c}$. To relieve the burden of feeding back FCSI information, distance- and effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ -based ordering strategies are considered in the literature. In this work, we propose a new ADI-based ordering strategy that uses ADI-based channel quality $\check{\zeta}$ in (4.15) and orders the users according to (4.16). In the first place, we consider the 2-user AD MIMO-NOMA scheme with fixed PA (FPA) [131], where a coefficient $0 \leq \psi \leq \frac{1}{2}$ controls the amount of power associated to each one of the pairing users such as $\gamma_{1,c} = \psi$ and $\gamma_{2,c} = 1 - \psi$. The proposed PA will be studied in Section 4.6.2.2

As seen in Fig. 4.4, and not surprisingly, the distance- and the effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ -based ordering strategies have very similar performance even in a congested cell. This is because $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ is inversely proportional to the distance, and thus the users will be sorted similarly from strong to weak users with the ascending order of distances or the descending order of effective channel gains. Interestingly, the proposed ordering strategy offers slightly better spectral efficiency than these two strategies. In fact, unlike the distance- or the effective channel gain-based ordering strategies, the proposed one takes into account spatial inter-cluster interference. In addition, the performance gap between the proposed ADI- and the optimal FCSI-based ordering strategies is fairly small, around 2% loss when 45 users are served simultaneously. With only the angular information feedback, i.e., with $2 \times M = 64$ fewer coefficients per channel estimation as shown in Table 4.2, the ADI-based ordering strategy provides a near-optimal sum-rate performance w.r.t. the optimal FCSI-based ordering strategy. Consequently, it is safe to conclude that the proposed ADI-based ordering strategy can be a viable solution in a directional mmWave

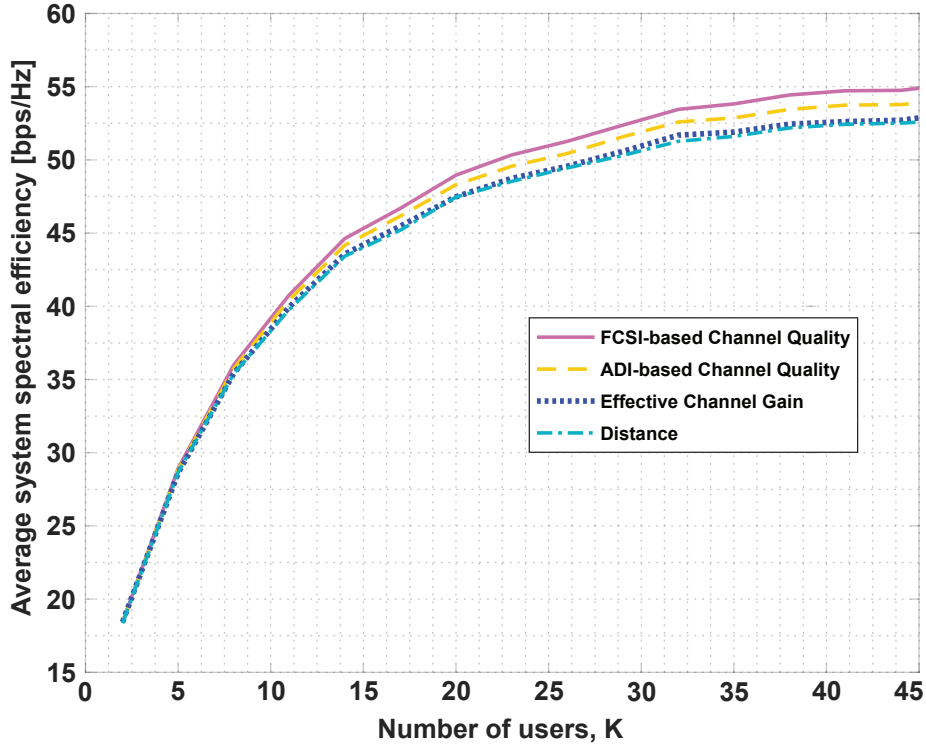


Figure 4.4: Average system spectral efficiency of 2-user AD MIMO-NOMA for different ordering strategies using distance, effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$, ADI-based or FCSI-based channel quality with FPA ($\psi = 0.4$).

environment.

4.6.2.2 Performance of power allocation technique

In this section, we aim to assess the intra-cluster PA proposed in Section 4.5.3, denoted by KKT-PA, for the 2-user AD MIMO-NOMA scheme by comparing it with FPA and fractional transmit PA (FTPA) [131]. In addition, through numerical simulations, we answer the question: does KKT-PA with ADI NOMA make an efficient SIC possible?

FTPA is similar to the power control used in the LTE uplink. FTPA is a channel dependent PA technique that dynamically allocates powers by taking into account the channel gains of users [131] as given by (4.41) and (4.42) for 2-user AD MIMO-NOMA with FCSI or ADI NOMA.

$$\gamma_{1,c}^{\text{type}} = \frac{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau}}{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau} + \left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}, \quad (4.41)$$

$$\gamma_{2,c}^{\text{type}} = \frac{\left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau} + \left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}, \quad (4.42)$$

The decay factor $0 \leq \tau \leq 1$ controls the power associated to users in NOMA cluster. ψ and τ are kept constant over all the NOMA clusters.

Fig. 4.5 and Fig. 4.6 show the average system spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA with $\psi \in \{0.1, 0.3, 0.5\}$ and FTPA with $\tau \in \{0, 1\}$ are applied, with FCSI and ADI NOMA, respectively.

As seen in Fig. 4.5 and Fig. 4.6, the average system spectral efficiency of the 2-user AD MIMO-NOMA scheme using FPA grows as ψ increases with either FCSI or ADI NOMA. This result is consistent with the analysis in Appendix A, which indicates that the sum-throughput is an increasing function of $\gamma_{1,c}$. In other words, increasing ψ increases the power allocated to the strong user, which in turn enhances the sum-throughput at the expense of decreasing the power allocated to the weak user. Using FTPA, the average system spectral efficiency decreases as τ increases. This is because more power is allocated to the weak user. When $\tau = 0$ and $\psi = 0.5$, FTPA and FPA are equivalent to equal transmit PA among users, that is the optimal solution in the absence of the efficient SIC constraint. However, KKT-PA involves the optimal solution $\gamma_{1,c}^{\star \text{type}} = \frac{1}{2} - \frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$ under the SIC constraints. From Fig. 4.5 and Fig. 4.6, it can be seen that the gap between KKT-PA and FPA with $\psi = 0.5$ in terms of spectral efficiency is negligible. More importantly, the dynamic KKT-PA maximizes the sum-throughput under the efficient SIC constraints and can be easily applied to any cluster size, unlike FPA and FTPA.

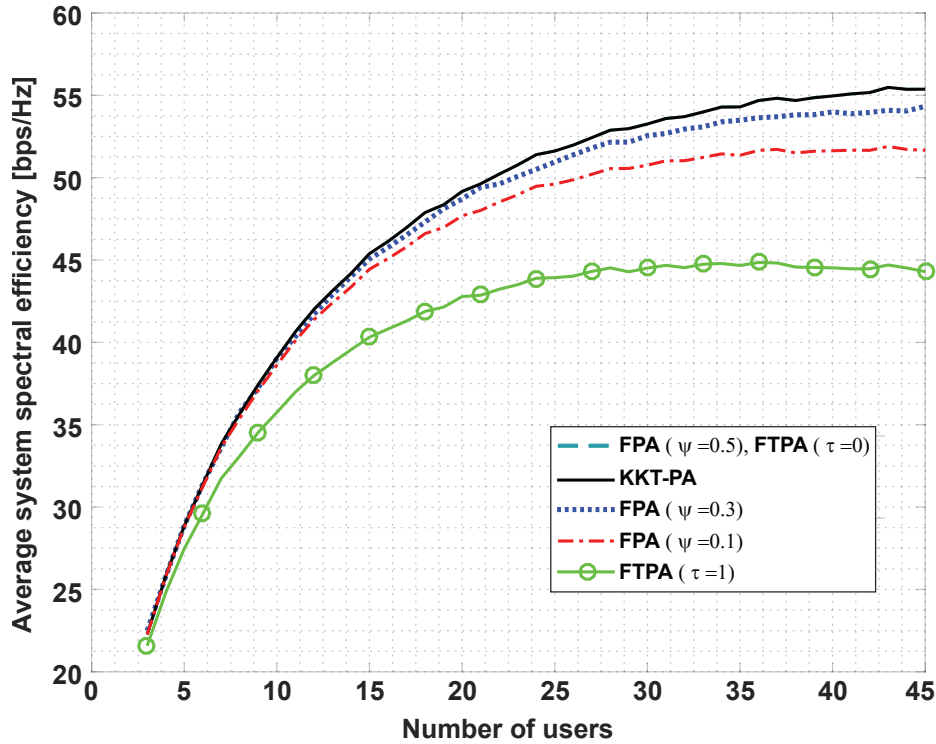


Figure 4.5: Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with FCSI NOMA.

Using KKT conditions, the closed-form solutions $\gamma_c^{\star \text{FCSI}}$ and $\gamma_c^{\star \text{ADI}}$ of $(\mathcal{P}_3)$ always satisfy that $\delta(\gamma_c^{\star \text{FCSI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}$ and $\delta(\gamma_c^{\star \text{ADI}}, \zeta_{l-1,c}^{\text{ADI}}) \geq P_{\min}$, respectively, as seen in (4.28b). However, we need to examine if the intra-cluster PA $\gamma_c^{\star \text{ADI}}$ derived from the ADI-based SIC constraint

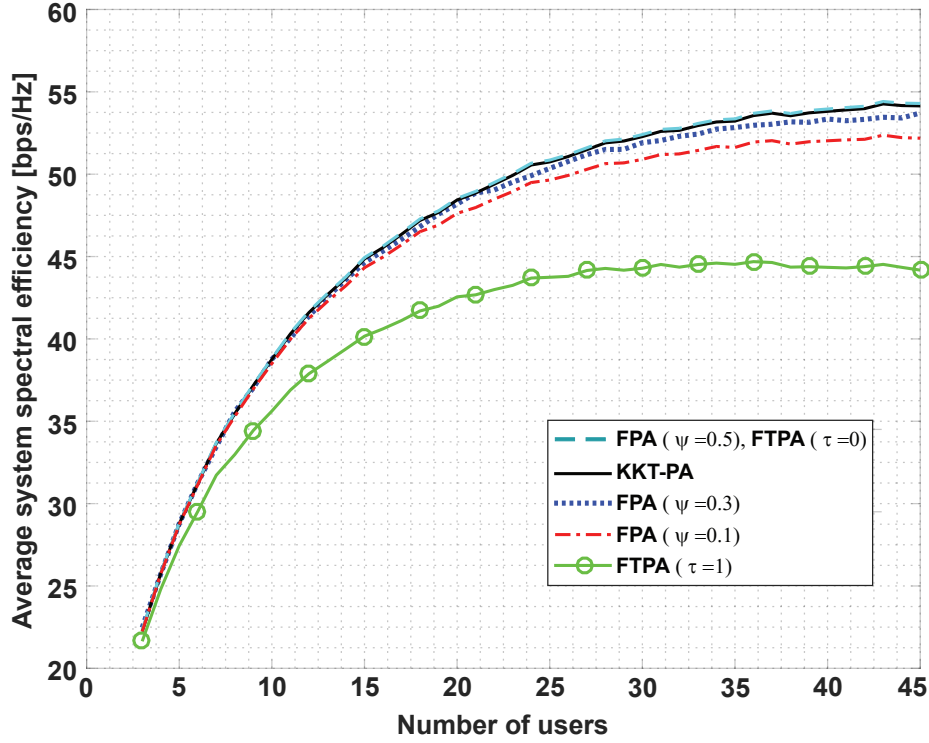


Figure 4.6: Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with ADI NOMA.

$\delta(\gamma_c^{\star\text{ADI}}, \zeta_{l-1,c}^{\text{ADI}}) \geq P_{\min}$ performs an efficient SIC in reality, i.e., if $\gamma_c^{\star\text{ADI}}$ satisfies that $\delta(\gamma_c^{\star\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}$. To this end, we calculate the probability Υ of efficient SIC among all the $C_n(K_c-1)$ users applying SIC when the KKT-PA with ADI NOMA is applied, such that

$$\Upsilon = P(\cup_{c=1}^{C_n} \cup_{l=2}^{K_c} \delta(\gamma_c^{\star\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}) = \sum_{c=1}^{C_n} \sum_{l=2}^{K_c} P(\delta(\gamma_c^{\star\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}). \quad (4.43)$$

We evaluate Υ in % for both 2-user and multi-user AD MIMO-NOMA with ADI NOMA when a different number of users, $K \in \{10, 20, 30, 40\}$, is served. $\Upsilon\%$ is averaged over 1000 trials. The numerical results provide that the average $\Upsilon\%$ is equal to 100% for both schemes and $\forall K \in \{10, 20, 30, 40\}$, which means that all users carefully apply efficient SIC. Consequently, the simplifying assumption used to design the angular-based NOMA scheme does not affect the effectiveness of SIC in highly directional channels.

4.6.2.3 Average spectral efficiency versus total number of users

To illustrate the advantages of the proposed multi-user and 2-user AD MIMO-NOMA schemes with either FCSI or ADI NOMA, we adopt two baseline schemes in our simulations. For baseline 1, C-AD-DBF scheme in [94] is considered, where the BS generates K directive beams, each one is steered in the spatial direction of the intended user. Note that C-AD-DBF can not support more than M users. For baseline 2, labeled as AD MIMO-OMA, ADI-based digital beamforming is associated with OMA scheme. For a fair comparison, the users are grouped into two categories of clusters, namely SU and 2-user clusters according to the same β -UC algorithm as in 2-user AD

MIMO-NOMA [99]. In 2-user OMA clusters, each user is served in its dedicated fraction of the orthogonal resource, e.g., time or frequency. Assume an equal resource allocation among users in the OMA cluster. Compared to C-AD-DBF, this scheme prevents users with a high level of spatial interference from getting a rate close to zero by grouping them in the same cluster and assigning them with orthogonal resources.

In Fig. 4.7, we investigate the average system spectral efficiency of the aforementioned schemes versus the total number of users for a spatial interference threshold of $\beta_0 = 0.5$. Performance sensibility w.r.t. the threshold is addressed in Section 4.6.4.

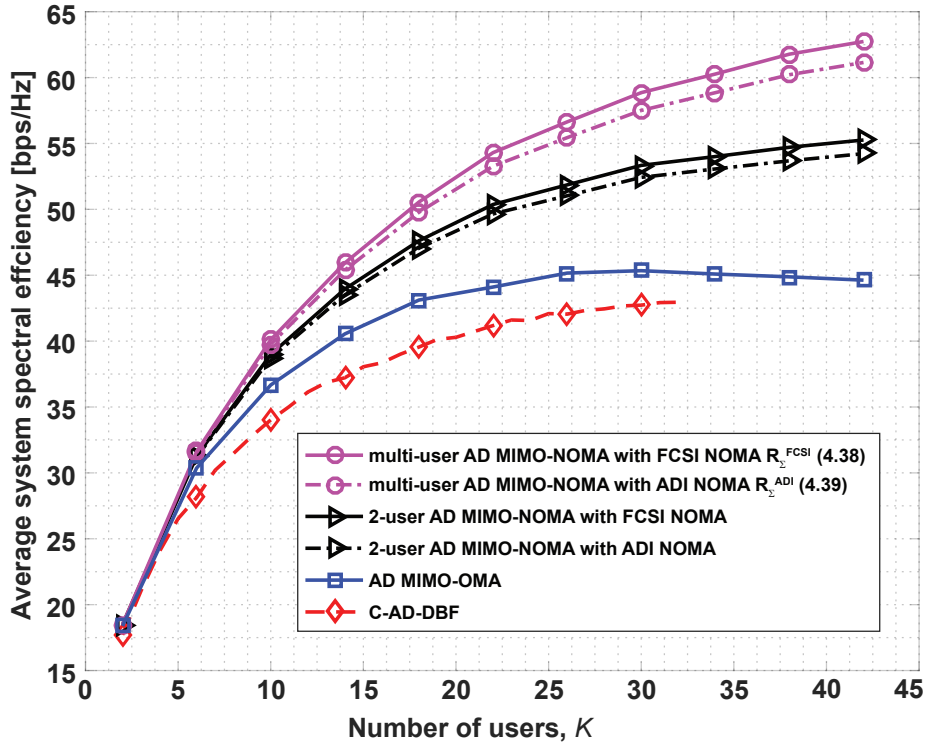


Figure 4.7: Average system spectral efficiency versus the total number of users, K , for $\beta_0 = 0.5$.

As observed from Fig. 4.7, attractive results on the performance of the proposed AD MIMO-NOMA with ADI NOMA are achieved, compared to FCSI NOMA. We see that the highest spectral efficiency gain $\frac{\Delta R_{\Sigma}}{B}$ (ΔR_{Σ} given by (4.40)), obtained for $K = 40$, of the proposed 2-user and multi-user AD MIMO-NOMA schemes using FCSI over that using ADI, is only 1.8% and 2.5%, respectively. Thus, the performance degradation due to the simplifying assumption, i.e., the ignorance of the NLoS paths and the high SNR regime, is very low. This observation verifies the interest of the proposed limited feedback NOMA protocol, including user ordering and intra-cluster power allocation, based only on ADI.

Moreover, Fig. 4.7 shows that ΔR_{Σ} increases with the number of users. However, since the number of NOMA clusters increases with the number of users, the simulation results in Fig. 4.7 are consistent with the analysis in Section 4.5.4 that ΔR_{Σ} rises with the number of NOMA clusters. Fig. 4.8 will analyze more precisely the variation of the number of users per cluster as K grows.

In Fig. 4.7, we observe that multiple access schemes such as NOMA and OMA improve the C-AD-DBF performance in terms of spectral efficiency. Indeed, all curves are above the C-AD-DBF one. This is due to the ability of the proposed β -UC algorithm to group users with high spatial interference and the potential of OMA and NOMA schemes to reduce this interference. Multi-user AD MIMO-NOMA with ADI NOMA achieves a spectral efficiency gain up to 36% w.r.t. C-AD-DBF, when K is close to M . More importantly, C-AD-DBF can not support simultaneously more than $M = 32$ users. Accordingly, NOMA is a promising technique to significantly boost the spectral efficiency of the C-AD-DBF system, particularly in a congested cell, at the expense of the additional complexity brought by the resource allocation (UC and PA) and the SIC decoding, at the transmitter and the receiver sides, respectively. A more detailed discussion of the complexity of our proposed scheme is given in Section 4.6.4.

When comparing OMA and NOMA protocols, we note that the average spectral efficiencies for the NOMA schemes increase as K grows, while it remains stable from $K = 25$ for OMA. Thus, NOMA schemes extensively outperform OMA for $K \geq 25$. Indeed, the NOMA schemes leverage the additional degrees of freedom brought by the power domain for multiple access, while OMA serves only one user in the same spectral or time resource block. This confirms the potential of AD MIMO-NOMA to exploit the multi-user diversity in a congested cell, unlike AD MIMO-OMA. However, using the same UC algorithm, the spectral efficiency gain of NOMA over OMA is achieved at the cost of greater decoding complexity at the receivers given by SIC and PA (in the case of equal repartition of PA for OMA clusters).

Finally, as expected, the proposed multi-user schemes outperform those with 2-user, especially in a congested cell. Indeed, as the number of users grows, the number of users per NOMA cluster also increases in the multi-user scheme to manage inter-user interference. In contrast, for the 2-user scheme, the number of users is fixed to two, and more beams, potentially in interference, are needed to address the growing number of users. This result will be discussed in the following Section 4.6.2.4.

In the rest of the work, we will only consider AD MIMO-NOMA with partial CSI, i.e., with ADI NOMA.

4.6.2.4 Number of users per NOMA cluster

Fig. 4.8 shows the impact of the proposed 2-user and multi-user β -UC algorithms on the scheduled clusters; their number, their types (SU or NOMA cluster) and the number of users in NOMA cluster, for a spatial interference threshold of $\beta_0 = 0.5$. Fig. 4.8a depicts the average number of clusters, using either 2-user or multi-user β -UC. Fig. 4.8b and Fig. 4.8c depict the cluster's percentage distribution per type for 2-user and multi-user β -UC, respectively, for $K \in \{5, 10, 20, 30, 40\}$.

As seen in Fig. 4.8a, the total number of clusters using 2-user β -UC is larger than when using multi-user β -UC, since multiple users $K_c \geq 2$ can be scheduled within the same cluster in the latter scheme. In Fig. 4.8b and Fig. 4.8c, we observe that, as the number of users increases, the number of 2- or multi-user clusters also grows to deal with the higher level of

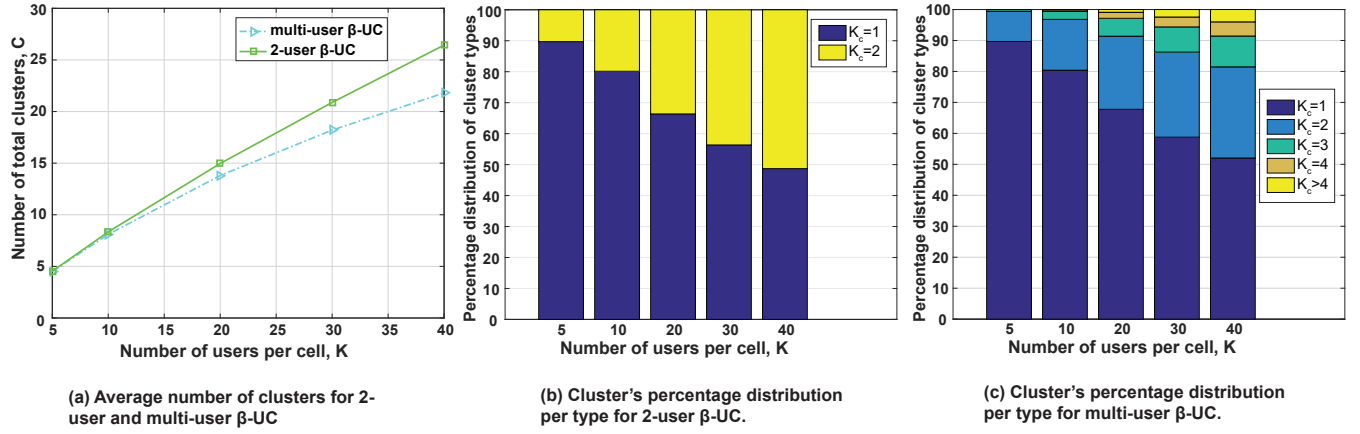


Figure 4.8: Average number of clusters (a), cluster's percentage distribution per type for 2-user (b) and multi-user AD MIMO-NOMA (c) versus the total number of users, K , for $\beta_0 = 0.5$.

spatial interference. Moreover, even if for $K = 5$, we get practically the same type of clusters, i.e., only 2-user, for both schemes, when K increases, multi-user clusters appear, with K_c equal to 3, 4 or greater than 4. Thanks to this additional degree of freedom, inter-cluster interference is reduced by passing from 2-user to multi-user scheme. Consequently, the multi-user scheme achieves a higher spectral efficiency w.r.t. the 2-user scheme, as shown in Fig. 4.7, especially when $K \geq 20$.

4.6.2.5 Average spectral efficiency versus spatial interference threshold

Fig. 4.9 investigates the impact of the spatial interference threshold β_0 on the system performance in terms of the average system spectral efficiency when $K \in \{10, 20, 40\}$ users are served simultaneously.

Firstly, we highlight that AD MIMO-NOMA with $\beta_0 = 1$ is equivalent to C-AD-DBF for $K \leq M$. The spectral efficiency curves of the proposed schemes change the slope around $\beta_0 = 0.21$. This observation is coherent with the value $\beta^{\text{FSL}} = 0.217$ given in Lemma 1. We recall that we consider the users as interferers, and we group them in the same NOMA cluster, only if the main lobe covers them within the β_0 -width with $\beta_0 \geq 0.217$. Instead, when $\beta_0 < 0.217$, users can interfere with secondary lobes. Based on this Lemma, the useful range is $0.217 \leq \beta_0 < 1$. In this range, we find that the two proposed schemes outperform C-AD-DBF (or AD MIMO-NOMA with $\beta_0 = 1$), as previously seen on Fig. 4.7.

In addition, the proposed multi-user scheme outperforms the 2-user one $\forall \beta_0 > 0.217$. In fact, the performance gap between these two schemes depends on two factors: (i) the total number of users, K and (ii) the spatial interference threshold, β_0 . This gap increases (resp. decreases) with growing K (resp. decreasing β_0). The increasing gap with K has already been observed in Fig. 4.7 and explained in section 4.6.2.4 by the additional degrees of freedom brought by the multi-user clusters. Varying β_0 captures the minimum level of inter-user interference between users in the same cluster. Then, when choosing lower β_0 satisfying Lemma 1, more users can be

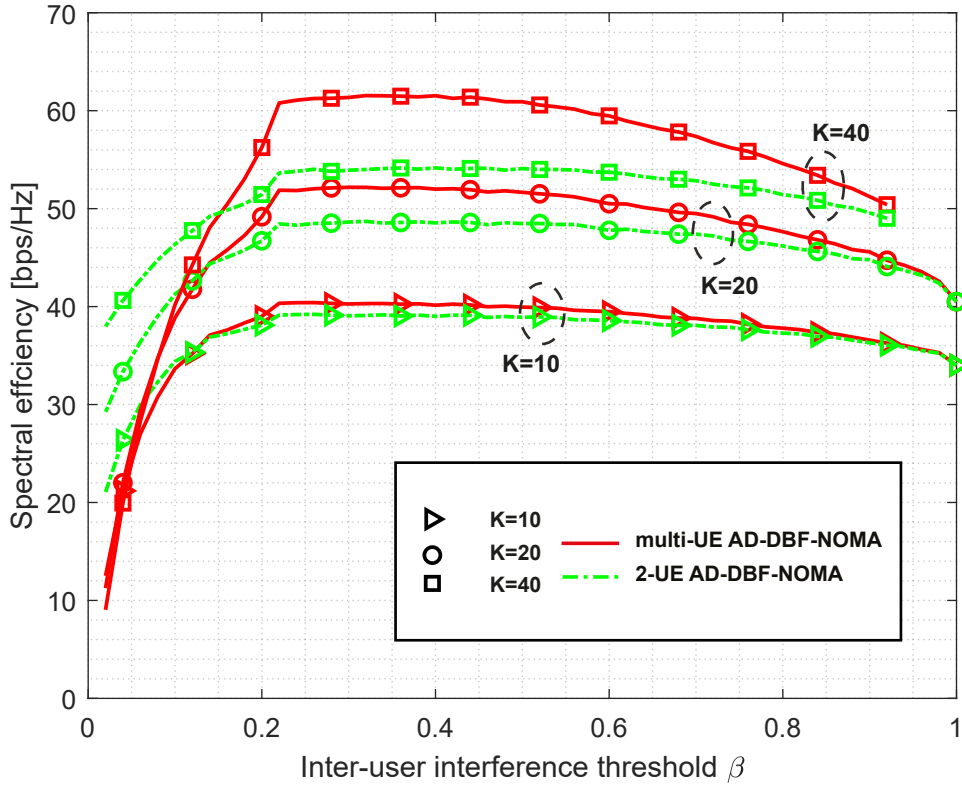


Figure 4.9: Average system spectral efficiency of AD MIMO-NOMA with ADI NOMA, for 2-user and multi-user clusters, when $K \in \{10, 20, 40\}$ users exist.

located in the same cluster. Thanks to the fewer clusters, inter-cluster interference is reduced, and the performance of the multi-user scheme is improved w.r.t. the 2-user one. Therefore, the lowest β_0 in the useful range produces the highest system spectral efficiency.

When comparing the multi-user AD MIMO-NOMA scheme with $\beta_0 = 0.217$ and $\beta_0 = 0.5$, the performance gap is negligible for $K = \{10, 20\}$. For K equal to 40, it is slightly in favor of $\beta_0 = 0.217$. Ultimately, as previously obtained, $\beta_0 = 0.217$ is the better choice. Moreover, for $\beta_0 \geq 0.217$, the choice of β_0 is not critical for the system performance.

4.6.3 3D AD MIMO-NOMA performance

Here, we evaluate the performance of the proposed AD MIMO-NOMA schemes when the BS adopts an $M_x \times M_z$ -URA array and serves $K \leq 45 < M$ users. Moreover, we investigate AD MIMO-NOMA from the energy efficiency (EE) perspective and verify the effectiveness of our schemes in improving the EE without deploying additional antennas and thereby with reduced power consumption.

4.6.3.1 Average spectral and energy efficiencies

In general, the spatial resolution to distinguish users depends on the width of the beam, i.e., the number of antennas in azimuth and elevation. The more antennas at the BS, the narrower the

beam is, which increases the received power and reduces interference between users, resulting in high SE. However, implementing massive antennas is not feasible due to high power consumption and transceiver complexity. To increase the SE and enable multi-user communication even in a congested cell, we propose to adopt NOMA as an alternative and innovative solution instead of increasing the number of antennas. Therefore, in this section, we compare the SE and EE performance of the 2-user and multi-user AD MIMO-NOMA schemes with C-AD-DBF, which uses more transmitting antennas.

Theoretically, the SE is not related to the power consumption. For m-MIMO systems in the limit of an infinite number of antennas, the power consumed at the RF circuits is a real challenge. Due to environmental and economic interests, the EE is also a key performance metric used for cellular networks and is defined as the ratio of the SE [bps] to the total power PC consumed at the BS and is given as follows [32]

$$EE = \frac{B \sum_{k=1}^K R_k}{PC} \quad [bps/J]. \quad (4.44)$$

The PC is calculated based on the effective transmit power ETP and the circuit power CP required for running the cellular systems [32]

$$PC = ETP + CP \quad [W]. \quad (4.45)$$

The ETP is given by $ETP = \frac{1}{\rho} P_e$, where $0 \leq \rho \leq 1$ is a constant modeling the inefficiency of power amplifiers. The CP is given by $CP = P_0 + MP_a$, where P_0 is the basic power consumed at the BS regardless of the number of antennas and P_a is the constant power consumption per antenna independent of the transmitted power (each antenna is assigned with an RF chain that contains many components; e.g., power amplifier, ADCs, DACs, local oscillators, filters, in-phase/quadrature (I/Q), etc.). Here, we set the power consumption parameters as in [32] as follows $\rho = 0.4$, $P_a = 1W$ and $P_0 = 10W$.

Fig. 4.10a and Fig. 4.10b depict the SE and the EE versus the number of users, respectively, for AD MIMO-NOMA using URA with $M_z = 2$, and for C-AD-DBF using URA with different values of M_z ($M_z = \{2, 4, 5, 6\}$) in a rural environment. Note that the different AD MIMO-NOMA and C-AD-DBF schemes have the same number of horizontal antennas, i.e., $M_x = 32$ and different number of vertical antennas, M_z .

Using 96 (resp. 128) less of antennas, the proposed 2-user (resp. multi-user) AD MIMO-NOMA scheme with $M_z = 2$ have almost the same performance of C-AD-DBF with $M_x = 32$ and $M_z = 5$ (resp. $M_x = 32$ and $M_z = 6$), as shown in Fig. 4.10a. Indeed, the additional antennas increase the BF gain and thus reduce inter-user interference. In contrast, AD MIMO-NOMA can improve the SE without deploying additional antennas by multiplexing the users with high inter-user interference in the power domain. Therefore, it's obvious that the gain EE of the system using 2-user (resp. multi-user) AD MIMO-NOMA is up to 129% (resp. 170%) against that using C-AD-DBF with $M_z = 5$ (resp. $M_z = 6$), as seen in Fig. 4.10b. Accordingly, AD MIMO-NOMA is an alternative solution to address the massive connectivity without significantly increasing

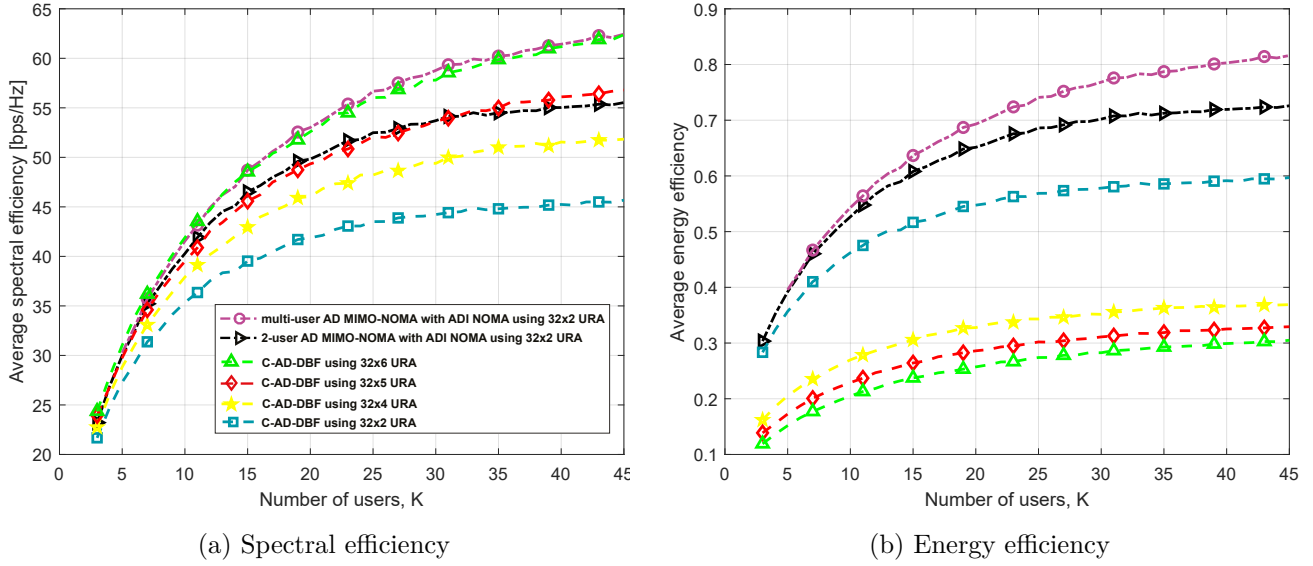


Figure 4.10: Average system spectral and energy efficiencies for 2-user and multi-user AD MIMO-NOMA using URA with $M_x = 32$ and $M_z = 2$ and for C-AD-DBF using URA with $M_x = 32$ and $M_z = \{2, 4, 5, 6\}$ in a rural environment. Both figures use the same legend presented in Fig. 4.10a.

the number of antennas.

4.6.4 AD MIMO-NOMA complexity

As already stated, the computational complexity of the proposed MIMO-NOMA schemes is brought by the resource allocation (UC and PA) at the BS and the SIC decoding at the receiver side. We discuss each of these independently for the 2-user and the multi-user schemes.

The complexity orders of both UC algorithms are given by (4.24). For 2-user β -UC, i.e., $K_c = 2$ ($\forall c \in C_n$) and $\lfloor \frac{K_c}{2} \rfloor = 1$. Therefore, the complexity order can be rewritten as $\mathcal{O}\left(\sum_{c=1}^{C_n} \mathbf{card}(\mathcal{F}_{1,c})\right)$, with $\mathbf{card}(\mathcal{F}_{1,c}) < (K-1)K \forall c \in C_n$. Now, we compare the complexity order of UC for both schemes by calculating the corresponding average of $\sum_{c=1}^{C_n} \sum_{p=1}^{\lfloor \frac{K_c}{2} \rfloor} \mathbf{card}(\mathcal{F}_{p,c})$ in Table 4.3. In addition, we evaluate in Table 4.3 the number of NOMA clusters C_n^{2-UE} and C_n^{m-UE} for 2-user and multi-user schemes, respectively. As seen in the Table, the multi-user β -UC algorithm has a lower complexity than the 2-user one, and the gap in the complexity between multi-user and 2-user schemes increases as K grows. This is because more NOMA clusters are scheduled in the 2-user scheme, as shown in Table 4.3. And on the other hand, because $\mathbf{card}(\mathcal{F}_{p,c})$ of the multi-user β -UC is decreasing faster than $\mathbf{card}(\mathcal{F}_{1,c})$ of the 2-user β -UC after each iteration due to the ability of multi-user β -UC to group multiple users in the same NOMA cluster.

The complexity order of the SIC decoding is linear with the number of SIC users³, i.e., $\mathcal{O}(\sum_{c=1}^{C_n} (K_c - 1))$. In addition, the complexity order of the PA method is linear with the number of all NOMA users, i.e., $\mathcal{O}(\sum_{c=1}^{C_n} K_c)$. Note that these expressions are the same for the multi-user and the 2-user schemes, but with different numbers of NOMA clusters, C_n , and different NOMA

³SIC users are the NOMA users applying SIC.

	2-user $\sum_{c=1}^{C_n} \text{card}(\mathcal{F}_{1,c})$	multi-user $\sum_{c=1}^{C_n} \sum_{p=1}^{\lfloor \frac{K_c}{2} \rfloor} \text{card}(\mathcal{F}_{p,c})$	C_n^{2-UE}	C_n^{m-UE}
$K = 10$	5.7	5.7	1.6	1.6
$K = 20$	35.6	33	4.9	4.4
$K = 30$	112.3	95.7	9.5	7.6
$K = 40$	258	201	13.6	10.2

Table 4.3: The complexity of 2-user and multi-user β -UC algorithms and the number of NOMA clusters of both schemes, when the BS serves $K \in \{10, 20, 30, 40\}$ users.

cluster sizes, K_c . Thus, in Table 4.4, to compare PA complexity orders (resp. SIC complexity orders), we calculate the relative percentage of NOMA users (resp. SIC users) of the multi-user scheme, as compared to the 2-user scheme.

	Relative percentage of NOMA users	Relative percentage of SIC users
$K = 20$	4.7%	19.3%
$K = 30$	5.9%	23.7%
$K = 40$	6.7%	25.8%

Table 4.4: Relative percentage of NOMA users and SIC users for the multi-user MIMO-NOMA scheme, as compared to the 2-user scheme, when the BS serves $K \in \{20, 30, 40\}$ users.

We observe that the relative percentage of NOMA users is low, even in a congested cell (e.g., 6.7% for $K = 40$). Indeed, the number of SU clusters is globally the same for both schemes and the remaining users are NOMA users. When looking at SIC complexity, however, with more users in each NOMA cluster (as shown in Fig. 4.8(c)), the relative percentage of SIC users is quite high, especially in a congested cell (e.g. 25.8% for $K = 40$). In other words, the gap in terms of SIC complexity increases as the number of users grows, unlike that of the PA method.

Finally, we conclude that SIC decoding highly contributes to the difference in complexity between the 2-user and the multi-user schemes.

4.7 Summary

In this chapter, we proposed 2-user and multi-user MIMO-NOMA transmission schemes, denoted as AD MIMO-NOMA, at mmWave frequencies using only the angular information to address the estimation and feedback problems of FCSI. This was done by considering the main features of mmWave channels, i.e., the high directionality and the potential blockage. We first derived a set of angular-based performance metrics, which participate in the design of user clustering, user order, and power allocation techniques. Specifically, the user clustering algorithms group the users with high spatial interference to perform NOMA, while the power allocation technique maximizes the system throughput. Based on numerical simulations, we obtained the following results:

- The proposed AD MIMO-NOMA scheme effectively improves the multi-user MIMO system throughput and guarantees the efficient SIC. For instance, multi-user AD MIMO-NOMA with ADI NOMA achieves a spectral efficiency gain up to 36% w.r.t. C-AD-DBF, when

K is close to M .

- The choice of the spatial interference threshold, β_0 , is critical for the system performance. To prevent the severe beam misalignment and for a good performance, β_0 must be larger than 0.217, which is the value of the FSL level.
- NOMA outperforms OMA to exploit the multi-user diversity in a congested cell.
- AD MIMO-NOMA is an alternative solution to address the massive connectivity without significantly increasing the number of antennas. For instance, using 128 less of antennas, the proposed multi-user AD MIMO-NOMA scheme with $M_z = 2$ have almost the same performance of C-AD-DBF with $M_x = 32$ and $M_z = 6$. Accordingly, the EE of the system using multi-user AD MIMO-NOMA with $M_z = 2$ is 4 times greater than that using C-AD-DBF with $M_z = 6$.

However, DBF with LSAs poses many challenging issues, including hardware complexity, cost, and power consumption. In the next chapter, we will consider the issues of DBF, and we will adopt the HBF technique.

Chapter 5

Angle-Domain Hybrid Beamforming-based MmWave Massive MIMO-NOMA

5.1 Introduction

Our results in Chapter 4 have shown the significant interest of the proposed AD DBF-based MIMO-NOMA schemes in the context of mmWave m-MIMO systems, where tens of antennas are employed at the BS (e.g., 32, 64). This interest includes the significant spectral efficiency gain and the low channel estimation complexity and overhead. Interestingly, m-MIMO systems can adopt hundreds of antennas at the BS to further enhance the spectral efficiency. Therefore, in the current chapter, we increase the high beamforming gain by using hundreds of antennas (e.g., 128, 256) at the BS.

The fully DBF requires as many RF chains as transmit antennas and is not feasible with m-MIMO, especially in the context of mmWaves, due to its high energy consumption and unaffordable hardware complexity and cost [87, 132]. Alternatively, the use of low-resolution ADCs is considered a viable way in m-MIMO systems [133–135] to address this issue. In addition, mmWave hybrid schemes with limited RF chains have been proposed as an effective solution for the possible implementations of mmWave m-MIMO systems [132, 136]. HBF, including both fully- and sub-connected structures, strikes a balance between complexity, energy consumption, and system performance. Both separate the signal processing into: 1) a low-dimensional digital precoder in the baseband (realized by a few RF chains) and 2) a high-dimensional analog precoder in the RF band to increase the array gain [132, 136, 137]. In [138], the authors compare the two kinds of hybrid structures for indoor mmWave and sub-THz systems. While the fully-connected structure achieves higher spectral efficiency, the sub-connected one can perform higher energy efficiency [87].

Another critical issue of m-MIMO is the requirement of a massive CSI overhead, which increases the signaling duration and produces a more extended training sequence. As mentioned

earlier, by exploiting the main features of mmWave channels, i.e., high directionality and significant blockage, the user's AoD, w.r.t. the BS, is considered as a promising partial CSI [86] for mmWave m-MIMO systems.

Therefore, in the current chapter, we focus on the AD HBF-based m-MIMO-NOMA, where the users are clustered based on their AoD information w.r.t. the BS. Specifically, the conventional AD MIMO-NOMA scheme generates a single beam in the spatial direction of the NOMA cluster. This scheme is denoted here as SB-NOMA. With LSAAs, the beamwidth is very narrow, and thus using SB-NOMA, only the users having similar AoD can be served concurrently within the same cluster. Hence, due to the limited number of RF chains in mmWave hybrid systems, SB-NOMA with LSAAs can not manage the connectivity of all users, especially in a congested cell. Hence, such a scheme lacks the potentiality to exploit the multi-user diversity in HBF-based m-MIMO systems.

The main objective of the current chapter is to overcome the restriction incurred by the few RF chains in mmWave hybrid systems and propose HBF-based mmWave m-MIMO-NOMA schemes with only the knowledge of angular information. Part of the work presented in this chapter has appeared in the International ITG Workshop on Smart Antennas (WSA) 2021 [139]. Next, we give an overview of the related works on HBF-based m-MIMO-NOMA.

5.1.1 Related works

HBF-based m-MIMO-NOMA systems are extensively studied in the literature to reduce energy consumption and hardware complexity [108, 129, 140–146]. In [108], the authors design a new HBF-based mmWave beamspace MIMO-NOMA scheme to support users more than the number of RF chains. And, an iterative algorithm that resolves the power allocation optimization problem is designed for maximizing the system sum-rate. In addition, the authors in [140] apply NOMA with the fully-connected HBF, and they develop new HBF technique by modifying the conventional block diagonalization scheme. This was done to improve the achievable spectral efficiency by reducing co-channel interference. The authors in [129] propose a user clustering algorithm according to the users' channel correlation and then formulate a joint HBF and power allocation problem for maximizing the system sum-rate under a minimum rate constraint. Yet, this problem is non-convex. To solve it, they first apply an arbitrary fixed HBF and find the corresponding solution of the power allocation. Then, they fix the ABF technique and design a DBF technique that reduces inter-group interference by applying the approximate ZF technique. Finally, they solve the ABF problem using the constant-modulus constraint with a proposed boundary-compressed particle swarm optimization algorithm. In [141], the authors consider both the sub- and fully-connected structures in the RF stage and ZF in the digital baseband. And, they propose an iterative low-complex power allocation algorithm to maximize energy efficiency. To improve both spectrum- and energy-efficiencies, the authors in [142, 143] integrate HBF-based MIMO-NOMA with the simultaneous wireless information and power transfer. In addition, the authors in [144] propose an optimal analog beamformer with the aid of ZF in the baseband to maximize both the sum-rate and the energy efficiency. This was done for the two

kinds of hybrid structures operating under LoS and NLoS mmWave environments. In [145], the authors use the signal-to-leakage ratio (SLNR) for the first time as the performance index to tackle the issues of the resource optimization in HBF-based m-MIMO-NOMA. Specifically, a joint optimization of power allocation and HBF is formulated such that the minimum user SLNR is maximized to ensure fairness between different users.

However, the previous works adopt a high-complex digital precoder and require full CSI knowledge. To reduce the channel estimation complexity and overhead with LSAAs on the one hand, and on the other hand, the complexity of signal processing in the baseband, this chapter will focus on AD MIMO-NOMA. Most existing works on AD MIMO-NOMA using ABF, DBF, or HBF partition the users into groups according to their angle difference and form a single beam toward each group [25, 86, 116–118]. However, as mentioned before, the beam-width is very narrow with LSAAs, and thus, SB-NOMA fails to accommodate all the users in mmWave hybrid systems.

To exploit more DoFs and accommodate more users using limited RF chains, the authors in [25] propose a joint SB-NOMA and TDMA scheme. Specifically, they cluster the users within single user (SU) and SB-NOMA groups, in which the angles belong to a predefined set. Subsequently, they propose a group clustering algorithm that schedules these groups in the time domain, so that inter-group interference in each time slot is reduced. This was done by considering that inter-group interference can be significantly reduced by increasing the spatial direction distance between the groups. While the SB-NOMA-TDMA scheme [25] reduces the complexity and the cost by using HBF, TDMA requires high precision in time synchronization among users since mmWave communications usually provide a high symbol rate. Moreover, the simultaneous application of TDMA with NOMA imposes more signal processing and synchronization complexity at the BS and the receiver sides.

Recently, the authors in [147] discussed the general idea of the multi-beam (MB) NOMA scheme in mmWave hybrid systems, and the authors in [148] offer its implementation details. Specifically, they propose a beam splitting technique that divides the entire transmit array into different subarrays. Accordingly, the BS can generate, within the same RF chain, multiple analog beams toward multiple NOMA users with arbitrary AoDs. The authors in [148] show that MB-NOMA efficiently exploits the multi-user diversity, unlike SB-NOMA, by performing NOMA transmission even when the users have very separate AoDs. In [149], the authors designed a suboptimal two-stage resource allocation that maximizes the system sum-rate based on full CSI. In the first stage, they assume that only ABF is available and then propose a joint user grouping and antenna allocation algorithm that maximizes the conditional system sum-rate by leveraging the coalition formation game theory. In the second stage, they adopt ZF in the digital baseband. Subsequently, they formulate a non-convex power allocation optimization problem to maximize the system sum-rate subject to the QoS constraints. To solve this problem, a suboptimal solution is devised. However, the scheme in [149] requires full channel knowledge and both schemes [25, 149] are designed only for the ULA architecture.

5.1.2 Contributions

In this chapter, we consider a mmWave hybrid system with fully-connected structure and LSAAs. Besides, we consider both ULA and URA architectures. And, we aim to exploit more DoFs for the SB-NOMA scheme, especially in an overloaded scenario, i.e., when the number of SU and multi-user SB-NOMA groups is larger than that of RF chains, N_{RF} . To do so, we propose two different schemes. While the first exploits the time-domain resources by using TDMA, the other only adopts NOMA without applying additional MA techniques, such as TDMA. Interestingly, these schemes require only the knowledge of angular information.

The main contributions of this chapter can be summarized as follows

- We propose two different schemes, namely joint β -based SB-NOMA and TDMA, and joint AD SB- and MB-NOMA schemes, to offer more DoFs for the conventional SB-NOMA scheme in mmWave hybrid systems.
- In the first phase of both schemes, we apply the β -based multi-user SB-NOMA clustering algorithm proposed in Chapter 4 to partition the users within SU and multi-user SB-NOMA clusters.
- In the second phase of the joint β -based SB-NOMA and TDMA scheme, we propose a group clustering algorithm inspired by [25]. Contrary to this work that measures inter-group interference by the angular distance and only applies with ULA, we use the spatial interference metric β proposed in Chapter 4 and apply the designed scheme with any uniformly excited antenna array architecture such as URA, uniform circular array (UCA), etc. Indeed, β is built based on the array factor, which implies that its use gives an accurate calculation of inter-group interference and enables the application of such scheme with different uniformly excited antenna array architectures. Specifically, inter-group interference is minimized in each time slot such that the total number of groups in each time slot is lower than N_{RF} .
- To reduce the complexity of signal processing at both Tx and Rx when applying TDMA with NOMA, we take profit of the MB-NOMA concept [148,149]. To do so, we first extend the beam splitting technique proposed in [149] for ULA to deal also with URA. Since we are interested only in the ADI feedback, we investigate the spatial behavior of SB-NOMA, MB-NOMA, and SDMA schemes for a 2-user scenario against the angular distance and the spatial interference metric β . We find that when the users are located in the main lobe of the beam of each other, SB-NOMA has a significant sum-rate performance. Otherwise, i.e., when the angular distance is larger than the 3dB beamwidth, MB-NOMA outperforms SB-NOMA and has approximately a considerable and constant sum-rate.
- We then extend the results obtained in a 2-user scenario to the general case of multi-user scenarios. Unlike the previous works on AD HBF-based m-MIMO-NOMA, we leverage the potentiality of SB-NOMA when the users are very close to each other and the ability of MB-NOMA to accommodate more users with separated AoDs. Specifically, we propose

two joint SB and MB-NOMA schemes, where the difference lies in the second phase. While the first one only uses MB-NOMA in an overloaded scenario to improve the sum-rate, the second one exploits MB-NOMA to reduce the number of active RF chains, i.e., the energy consumption as much as possible. Contrary to the work presented in [149] that proposes a user clustering algorithm based on the full CSI, we design a low-complex MB-NOMA clustering algorithm in the second phase with only the knowledge of angular information.

The rest of this chapter is organized as follows. The system model of HBF-based MIMO-NOMA is presented in Section 5.2. The proposed joint β -based SB-NOMA and TDMA, and joint AD-MB-SB-NOMA schemes are presented in Section 5.3 and Section 5.4, respectively. The performance of the proposed schemes is evaluated in Section 5.5. Finally, a summary of this chapter is given in Section 5.6.

5.2 System model

We consider a downlink HBF-based mmWave m-MIMO-NOMA system, where the BS equipped with $M = M_x M_z \gg 1$ antennas and $N_{RF} \ll M$ RF chains serves $K \ll M$ single-antenna UEs. We adopt a hybrid fully-connected structure at the BS. In this work, we consider both 1D and 2D antenna arrays at the BS by adopting a ULA array along the x -axis and a URA array in the xoz plane, respectively. For a ULA array along the x -axis, $M_z = 1$ and $\phi = 0$, thus, $M = M_x$ and $\vec{\Theta} = (\theta, 0)$. In the conventional AD MIMO-NOMA scheme, the users are regrouped within SU and multi-user SB-NOMA clusters, according to their AoDs. A block diagram of the scheme is given in Fig. 5.1.

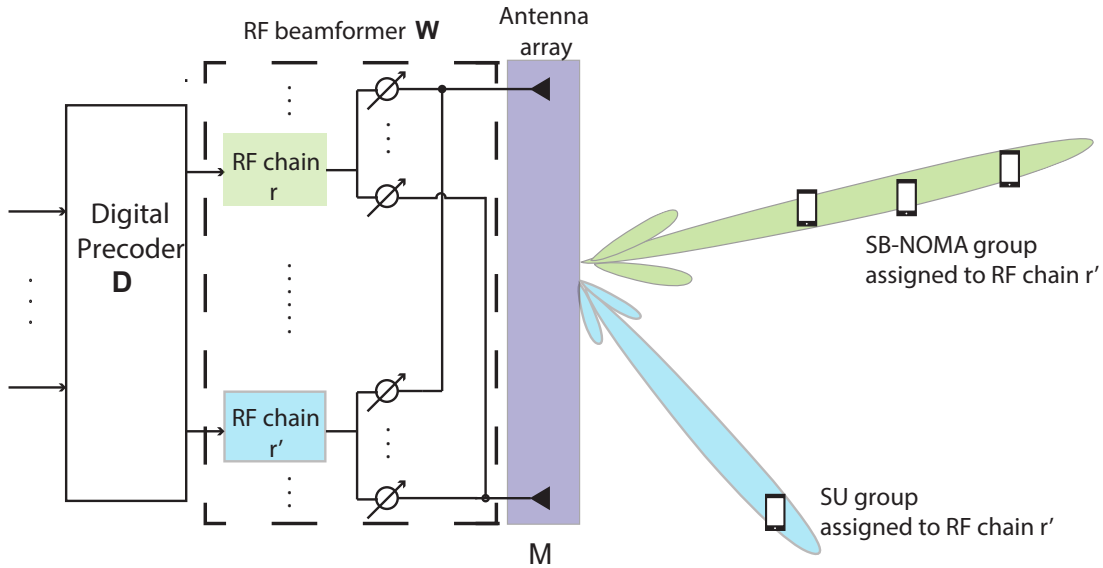


Figure 5.1: HBF-based SB-NOMA-MIMO scheme.

5.2.1 Hybrid beamforming design

Using HBF, the total number of groups G served simultaneously by the BS is restricted by the number of RF chains, i.e., $G \leq N_{RF}$. The overall downlink HBF matrix $\mathbf{F} \in \mathbb{C}^{M \times G}$ is constructed in two stages as follows

$$\mathbf{F} = \mathbf{W}\mathbf{D}, \quad (5.1)$$

where $\mathbf{D} = [\mathbf{d}_1 \cdots \mathbf{d}_G] \in \mathbb{C}^{G \times G}$ is the baseband digital component and $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_G] \in \mathbb{C}^{M \times G}$ is the RF analog component. Specifically, $\mathbf{d}_g$ and $\mathbf{w}_g$ denote respectively the digital and analog beamformers for the group g assigned to the g -th RF chain. In this work, we assume a pure analog mmWave system to reduce the complexity of the signal processing in the baseband, i.e., $\mathbf{D} = \mathbf{I}_G$. Thus, the normalization BF factor η given by $\eta = \frac{1}{\text{Tr}(\mathbf{F}^H \mathbf{F})}$ can be rewritten by $\eta = \frac{1}{\text{Tr}(\mathbf{W}^H \mathbf{W})}$.

5.2.2 HBF-based MIMO-NOMA system model

In the conventional AD HBF-based MIMO-NOMA scheme, the K users are clustered into G_{SU} SU groups and G_{SB} multi-user SB-NOMA groups. Denote $\mathcal{S}_g$, $g = 1, \dots, G$, as the set of users scheduled on the group g such that $\bigcup_{g=1}^G \mathcal{S}_g = \mathcal{K}$ and $\mathcal{S}_g \cap \mathcal{S}_{g'} = \emptyset$, $\forall g \neq g'$. We define the user scheduling variable u_k^g as follows

$$u_k^g = \begin{cases} 1, & \text{user } k \text{ belongs to group } g, \\ 0, & \text{otherwise.} \end{cases} \quad (5.2)$$

Note that each user will be served within one group, i.e., $\sum_{g=1}^G u_k^g = 1$, $\forall k \in \mathcal{K}$.

For both SU and multi-user SB-NOMA groups, the BS generates a single beam in the spatial direction $\vec{\Theta}_g$ of each group. Thereby, the analog beamformer of group g is given by

$$\mathbf{w}_g = a(\vec{\Theta}_g, M_x, M_z), \quad (5.3)$$

where $\vec{\Theta}_g$ is given similarly to (4.3) in Chapter 4 as follows

$$\vec{\Theta}_g = (\theta_g, \phi_g) = \begin{cases} (\theta_{1,k}, \phi_{1,k}), & \text{if SU with } \mathcal{S}_g = \{k\}, \\ \left(\theta_g = \frac{\min_{k \in \mathcal{S}_g} \{\theta_{1,k}\} + \max_{k \in \mathcal{S}_g} \{\theta_{1,k}\}}{2}, \phi_g = \frac{\min_{k \in \mathcal{S}_g} \{\phi_{1,k}\} + \max_{k \in \mathcal{S}_g} \{\phi_{1,k}\}}{2} \right), & \text{if SB-NOMA with } \text{card}(\mathcal{S}_g) \geq 2. \end{cases} \quad (5.4)$$

Denote by $\vec{\Theta}_{1,k} = (\theta_{1,k}, \phi_{1,k})$, the spatial direction of user k given by the AoD of the LoS path in user k channel.

We classify the users according to the angular-based user ordering strategy proposed in Section 4.3.1 that uses the angular-based channel quality $\check{\zeta}$ metric. Denote $\check{\zeta}_k^g$ as the angular-based channel quality of user k belonging to group g . Without loss of generality, we assume that the users are indexed in the descending order of their $\check{\zeta}$, i.e., $\check{\zeta}_k^g \geq \check{\zeta}_{k'}^{g'} \forall k \leq k'$. Therefore, the SIC decoding realizes the signal separation at the users side in the increasing order of $\check{\zeta}$.

Assuming a successful SIC decoding, the user k which belongs to group g receives the following signal y_k^g

$$y_k^g = \underbrace{\sqrt{\eta\gamma_{k,g}p_g}\mathbf{h}_k\mathbf{w}_gs_k}_{\text{desired signal}} + \underbrace{\sum_{k'=1}^{k-1} u_{k'}^g \sqrt{\eta\gamma_{k',g}p_g}\mathbf{h}_k\mathbf{w}_gs_{k'}}_{\text{intra-group interference}} + \underbrace{\sum_{\substack{g'=1 \\ g' \neq g}}^G \sum_{k'=1}^K u_{k'}^{g'} \sqrt{\eta\gamma_{k',g'}p_{g'}}\mathbf{h}_k\mathbf{w}_{g'}s_{k'}}_{\text{inter-group interference}} + z_k^g, \quad (5.5)$$

where s_k is the modulated signal relative to user k , p_g is the power allocated to group g such that $\sum_{g=1}^G p_g = P_e$ with P_e the total transmission power, $\gamma_{k,g}$ is the intra-group power allocation coefficient assigned to user k belonging to group g such that $\sum_{k=1}^K u_k^g \gamma_{k,g} = 1$ according to the NOMA principles, and $z_k^g \sim \mathcal{N}(0, \sigma_n^2)$ is the AWGN experienced at user k . Thereby, user k belonging to group g , has the following SINR_k^g , while decoding his own message

$$\text{SINR}_k^g = \frac{\eta\gamma_{k,g}p_g|\mathbf{h}_k\mathbf{w}_g|^2}{\sum_{k'=1}^{k-1} u_{k'}^g \eta\gamma_{k',g}p_g|\mathbf{h}_k\mathbf{w}_g|^2 + \sum_{\substack{g'=1 \\ g' \neq g}}^G \sum_{k'=1}^K u_{k'}^{g'} \eta\gamma_{k',g'}p_{g'}|\mathbf{h}_k\mathbf{w}_{g'}|^2 + \sigma_n^2}. \quad (5.6)$$

In (5.6), the first term in the denominator is residual intra-group interference after SIC, and the second one is inter-group interference. Considering a successful decoding at each user and no propagation error, user k belonging to group g achieves the following data rate R_k^g

$$R_k^g = \log_2(1 + \text{SINR}_k^g). \quad (5.7)$$

And, the total sum-rate is given by

$$R_T = \sum_{g=1}^G \sum_{k=1}^K u_k^g R_k^g. \quad (5.8)$$

5.2.3 Problem statement and proposed solutions

Due to the narrow beamwidth when using LSAAs, few users can be served within the SB-NOMA groups. In general, at most N_{RF} groups can be simultaneously connected to the BS via SDMA in mmWave hybrid systems. Thus, some users can not be served by the BS in a overloaded scenario, i.e., $G_{SU} + G_{SB} > N_{RF}$. The current chapter focuses on the limitation of SB-NOMA to offer more DoFs in mmWave hybrid systems with LSAA and designs two different schemes to confront this limitation. The first, namely joint β -based SB-NOMA and TDMA, leverages the time-domain resources and is inspired by [25]. Specifically, we cluster the obtained $G_{SU} + G_{SB}$ groups in different time slots such that spatial inter-group interference calculated with the β definition is reduced in each time slot. The other scheme, namely joint AD SB- and MB-NOMA scheme, inspired by [148], leverages the MB-NOMA framework, in which users with arbitrary AoDs can be served within the same group, i.e., by the same RF chain. Specifically, we cluster the remaining G_{SU} users within G'_{SU} SU and G_{MB} MB-NOMA groups, such that all the obtained groups can be served simultaneously by the BS without applying additional MA techniques

as TDMA, i.e., $G'_{SU} + G_{SB} + G_{MB} \leq N_{RF}$. In Section 5.3 and Section 5.4, we present the proposed joint β -based SB-NOMA and TDMA and the joint AD SB- and MB-NOMA schemes, respectively. Note that in this chapter, we consider the case when $2 \leq K \leq 2N_{RF}$, since we serve only two users within MB-NOMA clusters.

5.3 Joint β -based SB-NOMA and TDMA scheme

Algorithm 3: β -based group clustering in the time domain algorithm

Input: $\vec{\Theta}_g = (\theta_g, \phi_g)$, $\beta_{g,g'}$.

Output: N_{TS} , $\mathcal{G}_{TS}$

Initialization : $t = 0$, $\text{card}(\mathcal{G}_m^{\text{TS}}) = 0 \forall m \in \{1, \dots, N_{TS}\}$, $\mathcal{G}^{NA} = \{\text{group } g, g = 1, \dots, G\}$, $\mathcal{G}^{AV} = \{\mathcal{G}_1^{\text{TS}}, \dots, \mathcal{G}_{N_{TS}}^{\text{TS}}\}$.

1: Compute $\mathbf{B}$ as in (5.9).

Stage 1: Clustering of groups with high inter-group spatial interference in different time slots

2: **repeat**

3: Locate in $\mathbf{B}$ the 2 groups (e.g., groups o and r) having the largest spatial interference

4: **if** $o \in \mathcal{G}^{NA}$ **then**

5: $t = t + 1$, add group o to $\mathcal{G}_t^{\text{TS}}$, and remove group o from $\mathcal{G}^{NA}$.

6: **end if**

7: **if** $r \in \mathcal{G}^{NA}$ and $t < N_{TS}$ **then**

8: $t = t + 1$, add group r to $\mathcal{G}_t^{\text{TS}}$, and remove group r from $\mathcal{G}^{NA}$.

9: **end if**

10: Set $B(o, r) = -\infty$

11: **until** $t = N_{TS}$

Stage 2: Clustering of groups with low inter-group spatial interference in the same time slots

12: **repeat**

13: Set $t = t + 1$.

14: Select one arbitrary group g' from $\mathcal{G}^{NA}$

15: Locate, at each time-slot n , the group $\hat{i}_n$ having the highest spatial interference with the group g' , i.e., $\beta_{\hat{i}_n, g'} = \max_{i_n \in \mathcal{G}_n^{\text{TS}}} \beta_{i_n, g'}, \forall n \in \{1, \dots, \text{card}(\mathcal{G}_n^{\text{TS}})\}$

16: Find the time-slot $\check{n}$, in which the maximum interference with group g' obtained in Step 15 is minimized, i.e., $\beta_{\hat{i}_{\check{n}}, g'} = \min_{\{n=1, \dots, \text{card}(\mathcal{G}_n^{\text{TS}})\}} \beta_{\hat{i}_n, g'}$

17: Add group g' to $\mathcal{G}_{\check{n}}^{\text{TS}}$, and remove group g' from $\mathcal{G}^{NA}$.

18: **if** $\text{card}(\mathcal{G}_{\check{n}}^{\text{TS}}) \geq N_{RF}$ **then**

19: remove $\mathcal{G}_{\check{n}}^{\text{TS}}$ from $\mathcal{G}^{AV}$.

20: **end if**

21: **until** $\mathcal{G}^{NA} = \emptyset$.

The joint β -based SB-NOMA and TDMA is a two-phase scheme that applies both SB-NOMA and TDMA techniques to leverage more DoFs and accommodate several groups more than N_{RF} . In the first phase, the multi-user β -UC algorithm proposed in Chapter 4 partitions the K users within G_{SU} SU and G_{SB} multi-user SB-NOMA groups. Only in an overloaded scenario, i.e., $G_{SU} + G_{SB} > N_{RF}$, a second phase that exploits the time-domain resources is applied. In fact,

during each time slot, at most N_{RF} groups can be served concurrently. To this end, we need to design a group clustering algorithm that schedules the $G_{SU} + G_{SB} > N_{RF}$ groups in different time slots. In [25], the authors propose a two-stage group clustering algorithm that reduces inter-group interference in each time slot. Specifically, they use the spatial angular distance as a metric to measure the corresponding inter-group interference. And, they consider that inter-group interference can be significantly reduced by increasing the spatial direction distance. Moreover, the proposed algorithm is applied only with ULA. In this section, we have updated this group clustering such that inter-group interference measured by β is reduced in each time slot. Contrary to the work in [25], the β metric determines the level of spatial interference and includes both the angular distance and the beamwidth information, as mentioned in Chapter 4. Thereby, β is more accurate than the angular distance for calculating inter-group interference. In addition, the β metric is built based on the antenna array factor definition, and thereby the β -based group clustering algorithm can be applied to any antenna array architecture. Since $\beta_{g,g'} = \beta_{g',g}$, we first define the triangular matrix $\mathbf{B} \in \mathbb{C}^{G,G}$ describing spatial inter-group interference, such that the (g, g') -element in $\mathbf{B}$ is given by

$$\mathbf{B}(g, g') = \begin{cases} \beta_{g,g'}, & g < g', \\ -\infty, & \text{otherwise,} \end{cases} \quad (5.9)$$

where $\beta_{g,g'}$ is the spatial interference between group g and group g' , and is given by $\beta_{g,g'} = \frac{|\mathbf{a}^H(\tilde{\Theta}_g, M_x, M_z)\mathbf{a}(\tilde{\Theta}_{g'}, M_x, M_z)|}{M_x M_z}$.

Before detailing the proposed scheme, some notations and definitions are presented as follows

- N_{TS} is the total number of time slots and is given by $\lfloor \frac{G}{N_{RF}} \rfloor$ with $G = G_{SU} + G_{SB}$.
- $\mathcal{G}_m^{\text{TS}}$, $m = 1, \dots, N_{TS}$, is the group set including all the groups scheduled in the time slot m and is initialized by an empty set, i.e., $\text{card}(\mathcal{G}_m^{\text{TS}}) = 0$, $\forall m \in \{1, \dots, N_{TS}\}$.
- $\mathcal{G}^{NA}$ is the group set including all the groups not assigned yet to any time slot and is initialized by a set containing all the G groups, i.e., $\mathcal{G}^{NA} = \{\text{group } g, g = 1, \dots, G\}$.
- $\mathcal{G}^{AV}$ is the group set including all the available groups sets and is initialized by $\mathcal{G}^{AV} = \{\mathcal{G}_1^{\text{TS}}, \dots, \mathcal{G}_{N_{TS}}^{\text{TS}}\}$.

In mmWave hybrid systems, at most N_{RF} groups can be simultaneously scheduled in each time slot, i.e., $\text{card}(\mathcal{G}_m^{\text{TS}}) \leq N_{RF}$. Accordingly, $\mathcal{G}_m^{\text{TS}}$ is considered as an available group set if we can add at least one more group. In other words, if $\text{card}(\mathcal{G}_m^{\text{TS}}) < N_{RF}$, then $\mathcal{G}_m^{\text{TS}} \in \mathcal{G}^{AV}$.

In the first stage, the groups with high spatial interference β are scheduled in different time slots. Thus, N_{TS} groups are assigned to different time slots. Then, in the second stage, the remaining groups are clustered such that inter-group interference is reduced in each time slot. More details of the β -based group clustering in the time domain algorithm are given in Algorithm 3.

5.4 Joint angle-domain SB- and MB-NOMA schemes

In this section, we first present the basic concept of the MB-NOMA framework [148, 149] that generates multiple analog beams within the same RF chain using the beam splitting technique. Then, we investigate the spatial behavior of SB-NOMA, MB-NOMA, and SDMA schemes for a 2-user scenario since we are interested only in the ADI feedback. Finally, we extend the obtained results to a multi-user scenario, and we propose joint SB- and MB-NOMA schemes with different objectives. The first aims to increase the system spectral efficiency by adopting MB-NOMA only in an overloaded scenario, i.e., $G_{SU} + G_{SB} > N_{RF}$. The second one aims to reduce energy consumption as much as possible by applying MB-NOMA in both sparse and congested environments.

5.4.1 Multiple analog beams using beam splitting

The authors in [148, 149] propose an MB-NOMA scheme for mmWave hybrid systems to serve multiple users having arbitrary AoDs within the same RF chain. Specifically, they design a beam splitting technique that divides the entire antenna array into multiple subarrays to generate multiple analog beams. Interestingly, MB-NOMA can exploit the multi-user diversity more than SB-NOMA in mmWave hybrid systems. Indeed, using MB-NOMA, the number of users served simultaneously within the same NOMA cluster is not restricted by the users' AoD distribution as with SB-NOMA. In [148, 149], the beam splitting technique applies to the ULA architecture. However, since URA is more feasible with m-MIMO in practice [150], we extend this technique to deal with URA. Throughout this dissertation, we consider a 2-user MB-NOMA framework. In other words, we divide the entire antenna array connected with an RF chain into two subarrays with $M/2$ antennas to form two analog beams. And, we assume that both sub-arrays have the same size, i.e., the same number of antennas M_x^{sa} and M_z^{sa} along the x -axis and z -axis, respectively. Next, we separately present the beam splitting techniques with ULA and URA. To facilitate understanding the principle of this technique using either ULA or URA, we first rewrite the corresponding array steering vector in (2.3). Then, we consider user k and user k' belonging to the MB-NOMA group g , and we formulate the RF analog beamformer $\mathbf{w}_g$ assigned to the RF chain g performing beam splitting to generate two different beams toward each user.

5.4.1.1 Beam splitting with ULA

Using ULA with $M = M_x$ antennas along the x -axis, the array steering vector $\mathbf{a}(\vec{\Theta}, M_x, 1)$ corresponding to the angle $\vec{\Theta} = (\theta, 0)$ can be rewritten by

$$\mathbf{a}(\vec{\Theta}, M_x, 1) = \left[1, \dots, e^{j2\pi(M-1)\omega_x(\vec{\Theta})} \right]^T = \left[\mathbf{a}^T(\vec{\Theta}, M_x/2, 1), e^{j2\pi(M_x/2)\omega_x(\vec{\Theta})} \mathbf{a}^T(\vec{\Theta}, M_x/2, 1) \right]^T. \quad (5.10)$$

From (5.10), it is clear that $\mathbf{a}(\vec{\Theta}, M_x, 1)$ is a combination of the steering vectors of two subarrays with $M_x/2$ antennas separated by a phase shift term $e^{j2\pi(M_x/2)\omega_x(\vec{\Theta})}$.

Assume that the beam splitting technique divides the entire ULA array into two subarrays

with $M_x^{sa} = M_x/2$ and $M_z^{sa} = 1$. Therefore, the RF analog beamformer $\mathbf{w}_g$ assigned to the RF chain g performing beam splitting with two beams is given by [149]

$$\mathbf{w}_g = \left[\mathbf{a}^T(\vec{\Theta}_k, M_x^{sa}, M_z^{sa}), e^{j2\pi(M_x^{sa})\omega_x(\vec{\Theta}_k)} \mathbf{a}^T(\vec{\Theta}_{k'}, M_x^{sa}, M_z^{sa}) \right]^T. \quad (5.11)$$

From (5.11), each subarray generates a beam in the spatial direction of one user.

5.4.1.2 Beam splitting with URA

Similarly, using URA with M_x and M_z antennas along the x - and the z -axis, respectively, the array steering vector $\mathbf{a}(\vec{\Theta}, M_x, M_z) \in \mathbb{C}^{M_x M_z \times 1}$ corresponding to the angle $\vec{\Theta} = (\theta, \phi)$ can be rewritten by

$$\mathbf{a}(\vec{\Theta}, M_x, M_z) = \begin{bmatrix} \mathbf{a}^T(\vec{\Theta}, M_x, M_z/2) \\ e^{j2\pi(M_z/2)\omega_z(\vec{\Theta})} \mathbf{a}^T(\vec{\Theta}, M_x, M_z/2) \end{bmatrix} \quad (5.12)$$

Now, we extend the beam splitting technique to deal with URA. We divide vertically the antenna array into two subarrays, each one with $M_x^{sa} = M_x$ and $M_z^{sa} = M_z/2$ antenna elements along the x - and the z -axis, respectively, as shown in Fig. 5.2. And, the RF analog beamformer $\mathbf{w}_g$ associated to the RF chain g can be built based on the expression of $\mathbf{a}(\vec{\Theta}, M_x, M_z)$ in (5.12) as follows

$$\mathbf{w}_g = \begin{bmatrix} \mathbf{a}^T(\vec{\Theta}_k, M_x^{sa}, M_z^{sa}) \\ e^{j2\pi(M_z^{sa})\omega_z(\vec{\Theta}_k)} \mathbf{a}^T(\vec{\Theta}_{k'}, M_x^{sa}, M_z^{sa}) \end{bmatrix} \quad (5.13)$$

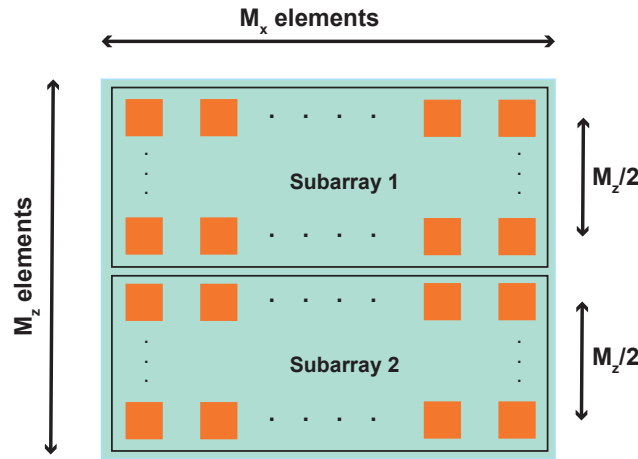


Figure 5.2: Vertically array splitting technique with URA.

From (5.13), it's clear that the first subarray can successfully form a beam toward the user k . Moreover, we multiply the steering vector of the second subarray by the same phase shift term $e^{j2\pi(M_z^{sa})\omega_z(\vec{\Theta}_k)}$. In other words, we add the same phase shift to all the antenna elements in the second subarray that have different phased weights to point a beam toward $\vec{\Theta}_{k'}$. Thereby, the second subarray can also successfully form a beam toward the user k' . It is worth noting that the beam splitting technique can be easily extended to support multiple users per RF chain,

i.e., generate multiple 2D or 3D analog beams with either ULA or URA.

5.4.1.3 Illustrative representations

To illustrate the generation of two analog beams via beam splitting, Fig. 5.3a and Fig. 5.3b plot the normalized array pattern responses for $\mathbf{w}_g$ in (5.11) and (5.13) with ULA and URA, respectively. Moreover, Fig. 5.3a and Fig. 5.3b plot the normalized array pattern responses of a single beam pointed at user k'' with ULA and URA, respectively. Compared to the single analog beam for user k'' , we can observe that the maximum magnitude of the main beam response is divided by two, i.e., the number of subarrays. This is because the two subarrays have the same dimensions. Moreover, since the number of antennas at each subarray is $M/2$, the beamwidth becomes wider for both beams pointed at users k and k' . In Section 5.4.2, we will show that the generation of multiple analog beams via beam splitting improves the system performance via serving more users on each RF chain when compared to SB-NOMA, which accommodates only the users having close directions.

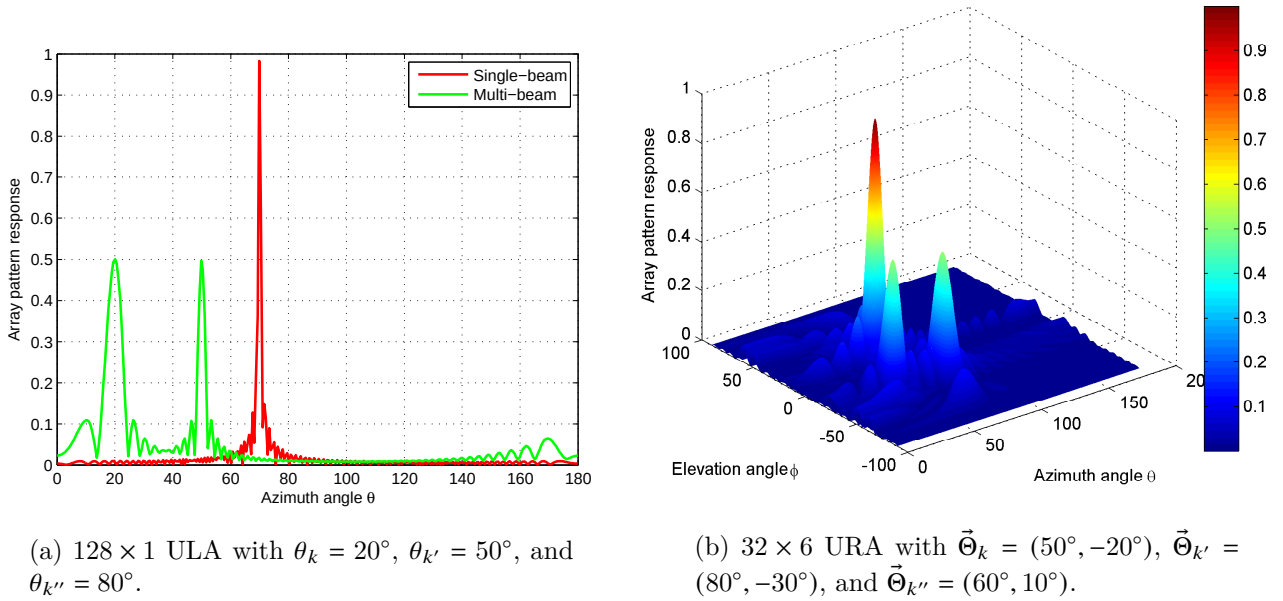


Figure 5.3: Antenna array pattern response for multiple beam pointed at users k and k' , and for a single beam pointed toward user k'' with (a) ULA and (b) URA.

5.4.2 2-user scenario: SB-NOMA Vs MB-NOMA Vs SDMA

The work presented in this thesis focuses on reducing the channel estimation overhead with LSAs by exploiting the mmWave channel properties and using only the ADI feedback. To this end, unlike the previous work on MB-NOMA that requires FCSI to perform the user clustering and does not leverage the potentiality of SB-NOMA when the users are close to each other, we need to investigate the spatial behavior of such schemes in mmWave channels. Hence, in this subsection, we derive the sum-rate of SB-NOMA, MB-NOMA, and SDMA schemes in a 2-user

scenario and study their spatial behavior against the angular distance and the spatial inter-user interference β . Throughout this subsection, we will answer the following questions. Does MB-NOMA have a considerable sum-rate? When SB-NOMA or MB-NOMA must be applied if we need to exploit more DoFs? How can we cluster the users in MB-NOMA groups using only angular information?

Here, we consider only two users served by the BS in a mono-path environment. Thereby, $\mathcal{K} = \{1, 2\}$ and only one group exists in either SB-NOMA or MB-NOMA scheme, i.e., $G = 1$. Assume that user 1 is the strong user. Next, we will give the system sum-rate using SDMA, SB-NOMA, or MB-NOMA. Then, we will analyze their performance and compare them.

5.4.2.1 System sum-rate

In the following, we give the system sum-rate when using either SDMA, SB-NOMA or MB-NOMA.

1. SDMA

When SDMA is applied, each user is served by a single beam. And, the $\text{SINR}_1^{\text{SDMA}}$ at user 1 is given by ($\text{SINR}_2^{\text{SDMA}}$ can be obtained by symmetry)

$$\text{SINR}_1^{\text{SDMA}} = \frac{\eta P_e |\mathbf{h}_1 \mathbf{a}(\vec{\Theta}_{1,1}, M_x, M_z)|^2}{\eta P_e |\mathbf{h}_1 \mathbf{a}(\vec{\Theta}_{1,2}, M_x, M_z)|^2 + \sigma_n^2}, \quad (5.14a)$$

$$\stackrel{(a)}{=} \frac{\eta P_e |\alpha_{1,1}|^2 |\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x, M_z) \mathbf{a}(\vec{\Theta}_{1,1}, M_x, M_z)|^2}{\eta p_2 |\alpha_{1,1}|^2 |\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x, M_z) \mathbf{a}(\vec{\Theta}_{1,2}, M_x, M_z)|^2 + \sigma_n^2}, \quad (5.14b)$$

$$\stackrel{(b)}{=} \frac{\rho \eta |\alpha_{1,1}|^2 M^2}{\rho \eta p_2 |\alpha_{1,1}|^2 |\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x, M_z) \mathbf{a}(\vec{\Theta}_{1,2}, M_x, M_z)|^2 + 1}, \quad (5.14c)$$

where (a) is obtained from the mono-path channel assumption, i.e., $\mathbf{h}_k = \alpha_{1,k} \mathbf{a}^H(\vec{\Theta}_{1,1}, M_x, M_z)$, and (b) is obtained since $|\mathbf{a}^H(\vec{\Theta}, M_x, M_z) \mathbf{a}(\vec{\Theta}, M_x, M_z)| = M_x M_z = M$ and by setting $\rho = P_e / \sigma_n^2$.

By setting $\beta_{1,2}(M_x, M_z)$ as follows

$$\beta_{1,2}(M_x, M_z) = \frac{|\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x, M_z) \mathbf{a}(\vec{\Theta}_{1,2}, M_x, M_z)|}{M_x M_z}, \quad (5.15)$$

and since η is given by $\eta = \frac{1}{2M}$ when applying SDMA, the system sum-rate R_T^{SDMA} can be expressed as follows

$$R_T^{\text{SDMA}} = \sum_{k=1}^2 \log_2 \left(1 + \frac{\rho |\alpha_{1,k}|^2 M}{\rho |\alpha_{1,k}|^2 \beta_{1,2}^2(M_x, M_y) M + 2} \right) \quad (5.16)$$

2. SB-NOMA

SB-NOMA forms a single beam between the two users, with an angle calculated as in (5.4). Thereby, $\text{SINR}_{1,g}^{\text{SB}}$ and $\text{SINR}_{2,g}^{\text{SB}}$ ($g = 1$) of the two users are respectively given by

$$\text{SINR}_{1,g}^{\text{SB}} = \frac{\eta\gamma_{1,g}p_g|\mathbf{h}_1\mathbf{w}_g|^2}{\sigma_n^2}, \quad (5.17a)$$

$$\stackrel{(a)}{=} \rho\eta\gamma_{1,g}|\alpha_{1,1}|^2|\mathbf{a}^H(\theta_{1,1}, M)\mathbf{a}(\theta_g, M)|^2, \quad (5.17b)$$

where (a) is obtained since one group exists, i.e., $p_g = P_e$.

$$\text{SINR}_{2,g}^{\text{SB}} = \frac{\eta(1 - \gamma_{1,g})p_g|\mathbf{h}_2\mathbf{w}_g|^2}{\gamma_{1,g}p_g|\mathbf{h}_2\mathbf{w}_g|^2 + \sigma_n^2}, \quad (5.18a)$$

$$= \frac{\rho\eta(1 - \gamma_{1,g})|\alpha_{1,2}|^2\left|\mathbf{a}^H(\vec{\Theta}_{1,2}, M_x, M_z)\mathbf{a}(\vec{\Theta}_g, M_x, M_z)\right|^2}{\rho\eta\gamma_{1,g}|\alpha_{1,2}|^2\left|\mathbf{a}^H(\vec{\Theta}_{1,2}, M_x, M_z)\mathbf{a}(\vec{\Theta}_g, M_x, M_z)\right|^2 + 1}. \quad (5.18b)$$

We set $\beta_{k,g}(M_x, M_z) = \frac{|\mathbf{a}^H(\vec{\Theta}_g, M_x, M_z)\mathbf{a}(\vec{\Theta}_{1,k}, M_x, M_z)|}{M_x M_z}$, and since η is given by $\eta = \frac{1}{M}$ when applying SB-NOMA, the system sum-rate R_T^{SB} can be expressed as follows

$$R_T^{\text{SB}} = \log_2 \left(1 + \rho\gamma_{1,g}|\alpha_{1,1}|^2\beta_{1,g}^2(M_x, M_z)M \right) + \log_2 \left(1 + \frac{\rho(1 - \gamma_{1,g})|\alpha_{1,2}|^2\beta_{2,g}^2(M_x, M_z)M}{\rho\gamma_{1,g}|\alpha_{1,2}|^2\beta_{2,g}^2(M_x, M_z)M + 1} \right) \quad (5.19)$$

3. MB-NOMA

MB-NOMA generates two beams within the same RF chain, each one is pointed in the AoD of one user by using the beam splitting technique. The analog beamformer $\mathbf{w}_g$ ($g = 1$) is given by (5.11) and (5.13) with ULA and URA, respectively. Denote by $\psi(\vec{\Theta}_k, M_x^{sa}, M_z^{sa})$ the additional phase shift experienced at the second subarray using either ULA or URA, and is given by

$$\psi(\vec{\Theta}_k, M_x^{sa}, M_z^{sa}) = \begin{cases} j2\pi(M_x^{sa})\omega_x(\vec{\Theta}_k) & \text{if ULA,} \\ j2\pi(M_z^{sa})\omega_z(\vec{\Theta}_k) & \text{if URA.} \end{cases} \quad (5.20)$$

Thus, using either ULA or URA, $\text{SINR}_{1,g}^{\text{MB}}$ and $\text{SINR}_{2,g}^{\text{MB}}$ of the two users are respectively given by

$$\text{SINR}_{1,g}^{\text{MB}} = \frac{\eta\gamma_{1,g}p_g|\mathbf{h}_1\mathbf{w}_g|^2}{\sigma_n^2}, \quad (5.21a)$$

$$= \rho\eta\gamma_{1,g}|\alpha_{1,1}|^2 \left| M/2 + \mathbf{a}^H(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa})\mathbf{a}(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa}) \right|^2. \quad (5.21b)$$

$$\text{SINR}_{2,g}^{\text{MB}} = \frac{\eta(1 - \gamma_{1,g})p_g|\mathbf{h}_2\mathbf{w}_g|^2}{\gamma_{1,g}p_g|\mathbf{h}_2\mathbf{w}_g|^2 + \sigma_n^2}, \quad (5.22a)$$

$$= \frac{\rho\eta(1 - \gamma_{1,g})|\alpha_{1,2}|^2 \left| \mathbf{a}^H(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa}) \mathbf{a}(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) + M/2e^{\Delta\psi} \right|^2}{\rho\eta\gamma_{1,g}|\alpha_{1,2}|^2 \left| \mathbf{a}^H(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa}) \mathbf{a}(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) + M/2e^{\Delta\psi} \right|^2 + 1}, \quad (5.22b)$$

where $\Delta\psi = \psi(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) - \psi(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa})$.

We set $\beta'_{1,2}(M_x^{sa}, M_z^{sa}) = \frac{\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) \mathbf{a}(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa})}{M_x^{sa} M_z^{sa}} = \frac{\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) \mathbf{a}(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa})}{M/2}$, and since η is given by $\eta = \frac{1}{M}$ when applying MB-NOMA, thus the system sum-rate R_T^{MB} can be expressed as follows

$$\begin{aligned} R_T^{\text{MB}} = & \log_2 \left(1 + \rho\gamma_{1,g}|\alpha_{1,1}|^2 \left| \frac{1}{2} \left(1 + \beta'_{1,2}(M_x, M_z/2) \right) \right|^2 M \right) + \\ & \log_2 \left(1 + \frac{\rho(1 - \gamma_{1,g})|\alpha_{1,2}|^2 \left| \frac{1}{2} \left(\beta'_{1,2}(M_x^{sa}, M_z^{sa}) + e^{\Delta\psi} \right) \right|^2 M}{\rho\gamma_{1,g}|\alpha_{1,2}|^2 \left| \frac{1}{2} \left(\beta'_{1,2}(M_x^{sa}, M_z^{sa}) + e^{\Delta\psi} \right) \right|^2 M + 1} \right) \end{aligned} \quad (5.23)$$

5.4.2.2 Performance analysis

For 128×1 ULA, Fig. 5.4a and Fig. 5.4b plot the sum-rate of the three schemes against respectively the angular distance $\Delta\theta$ and the spatial inter-user interference $\beta_{1,2}$ calculated in (5.15). The AoD $\theta_{1,1}$ of the first user is set to 10° , and we uniformly change the AoD $\theta_{1,2}$ of the second user, such that $0^\circ \leq \Delta\theta = \theta_{1,2} - \theta_{1,1} \leq 120^\circ$. For 16×16 URA, Fig. 5.5 and Fig. 5.6 plot the sum-rate against the angular distance $\Delta\vec{\Theta}$ and the spatial inter-user interference $\beta_{1,2}$, respectively. The AoD $\vec{\Theta}_{1,1}$ of the first user is set to $(30^\circ, -60^\circ)$, and we uniformly change the AoD $\vec{\Theta}_{1,2}$ of the second user, such that $0^\circ \leq \Delta\theta = \theta_{1,2} - \theta_{1,1} \leq 120^\circ$ and $0^\circ \leq \Delta\phi = \phi_{1,2} - \phi_{1,1} \leq 60^\circ$. Note that for ULA and URA, we set $\gamma_{1,g} = 0.2$, $\gamma_{2,g} = 0.8$, $Pe = 30\text{dBm}$ and $\sigma_n^2 = -101\text{dBm}$. For ULA, $|\alpha_{1,1}| = 8.6 \times 10^{-6}$ and $|\alpha_{1,2}| = 6.2 \times 10^{-6}$. For URA, $|\alpha_{1,1}| = 7.7 \times 10^{-6}$ and $|\alpha_{1,2}| = 2.9 \times 10^{-5}$.

From (5.16), we find that R_T^{SDMA} is a decreasing function w.r.t $\beta_{1,2}$. This also can be seen from Fig. 5.4b and Fig. 5.6 for ULA and URA, respectively. In other words, when the users are very close to each other, i.e., $\Delta\vec{\Theta} \rightarrow (0,0)$ and $\beta_{1,2} \rightarrow 1$, the system suffers from high inter-user interference and the SDMA performance is degraded. For instance, for a large number of transmit antennas but finite or in a high SNR regime, R_T^{SDMA} in (5.16)

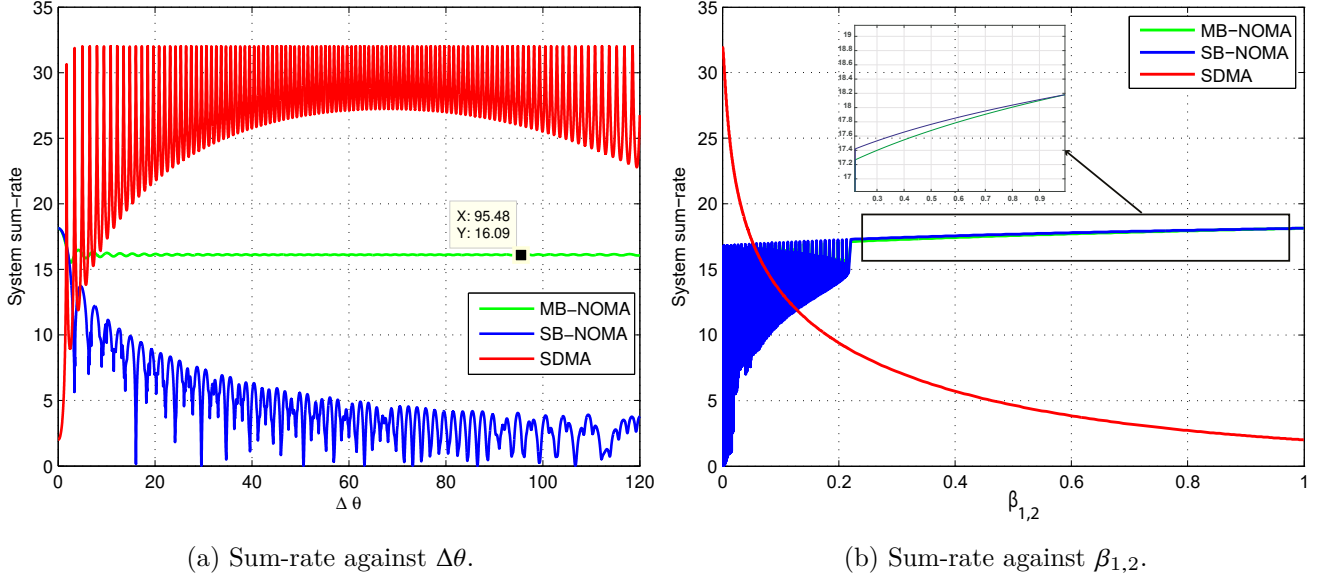


Figure 5.4: SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\Delta\theta$ and $\beta_{1,2}$ with 128×1 ULA.

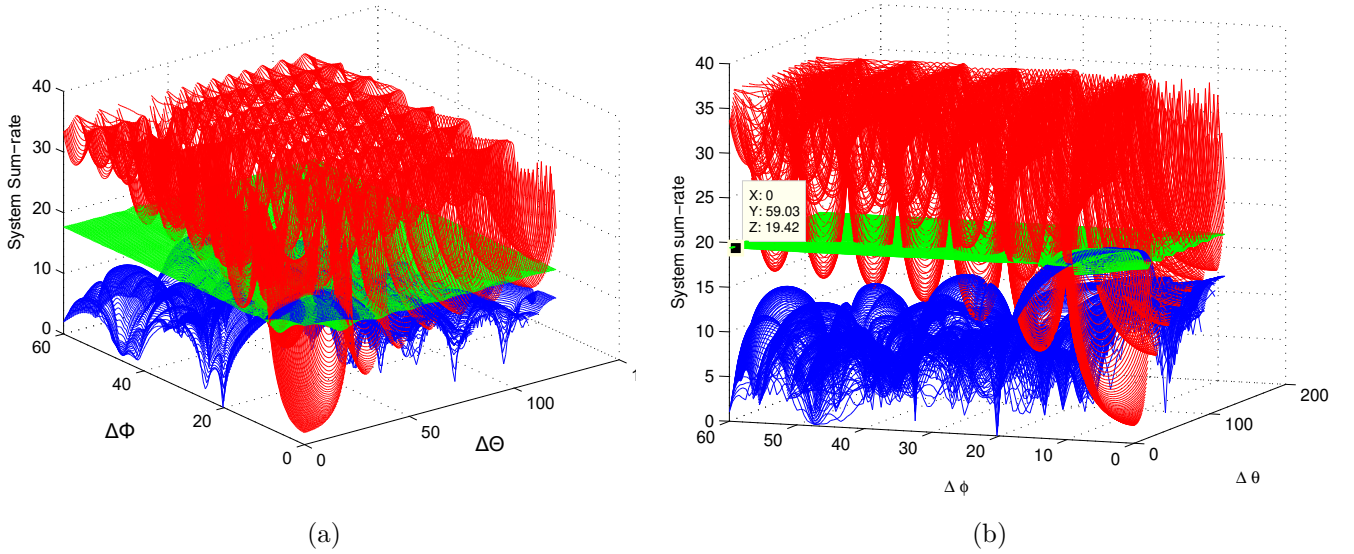


Figure 5.5: SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\Delta\theta$ and $\Delta\phi$ with 16×16 URA.

can be approximated as $R_T^{\text{SDMA}} \xrightarrow[\beta_{1,2} \rightarrow 1]{\Delta\vec{\theta} \rightarrow (0,0)} K = 2$. From Fig. 5.4, Fig. 5.5 and Fig. 5.6, we can see that $R_T^{\text{SDMA}} = 2$ when $\Delta\vec{\theta} = (0,0)$, i.e., $\beta_{1,2} = 1$, and that SDMA has a significant performance when the users are well separated in space. With large and finite number of transmit antennas, $|\mathbf{a}^H(\vec{\theta}_{1,1}, M_x, M_z)\mathbf{a}(\vec{\theta}_{1,2}, M_x, M_z)| \rightarrow 0$ when $\Delta\vec{\theta} \neq (0,0)$ [151]. Thereby,

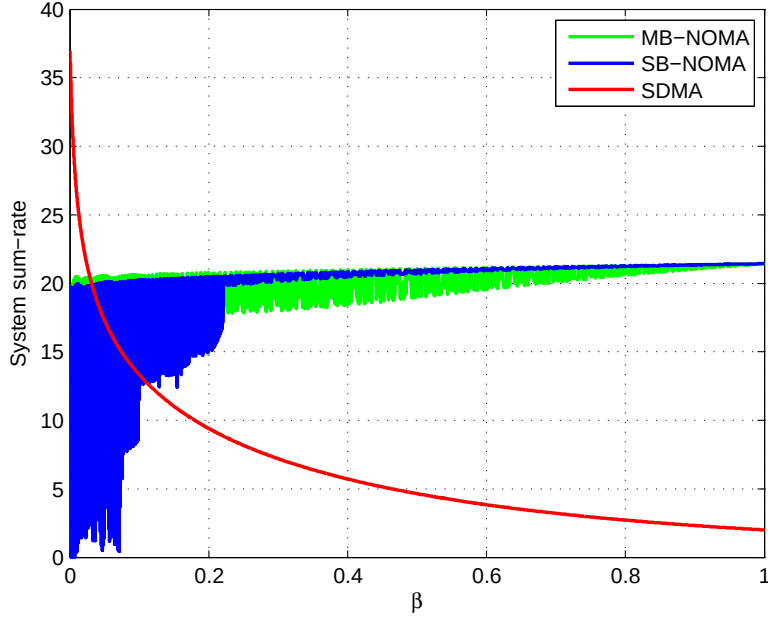


Figure 5.6: SB-NOMA versus MB-NOMA versus SDMA: Sum-rate against $\beta_{1,2}$ with 16×16 URA.

R_T^{SDMA} can be approximated to

$$R_T^{\text{SDMA}} \xrightarrow{M \gg 1} \sum_{k=1}^2 \log_2 \left(1 + \frac{\rho |\alpha_{1,k}|^2 M}{2} \right). \quad (5.24)$$

From (5.19), R_T^{SB} is restricted by the $\beta_{k,g}$ component. From Lemma 2 in Chapter 4, if user 2 is located in the main lobe of the beam pointed to user 1, i.e., $\Delta \vec{\Theta} \leq \vec{\Omega}_1^{\text{3dB}} = (\Omega_{x1}^{\text{3dB}}, \Omega_{z1}^{\text{3dB}})^1$ and $\beta_{1,2} \geq \beta^{\text{FSL}} = 0.217$, then both users will be located in the main lobe of the beam directed at $\vec{\Theta}_g$. Therefore, SB-NOMA works well and the system sum-rate is maximized. More the users are separated, more SB-NOMA degrades due to the beam misalignment. This is also observed in Fig. 5.4, Fig. 5.5 and Fig. 5.6. With large and finite number of transmit antennas, $|\mathbf{a}^H(\vec{\Theta}_g, M_x, M_z) \mathbf{a}(\vec{\Theta}_{1,k}, M_x, M_z)| \rightarrow 0$ for $\Delta \vec{\Theta} \neq (0, 0)$, thereby, $R_T^{\text{SB}} \xrightarrow{M \gg 1} 0$.

From (5.19), R_T^{MB} is restricted by the $\beta'_{1,2}(M_x^{sa}, M_z^{sa})$ component. When the two users have similar AoD, i.e., $\Delta \vec{\Theta} \rightarrow (0, 0)$, the analog beamformer $\mathbf{w}_g$ is degenerated to the single-beam case, and thereby $R_T^{\text{MB}} \rightarrow R_T^{\text{SB}}$, which is the highest sum-rate when using NOMA. From Fig. 5.4, Fig. 5.5 and Fig. 5.6, when $0 < \Delta \vec{\Theta} \leq \vec{\Omega}_1^{\text{3dB}} = (\Omega_{x1}^{\text{3dB}}, \Omega_{z1}^{\text{3dB}})$, SB-NOMA slightly outperforms MB-NOMA since the two beams generated by MB-NOMA scheme are superimposed but with small beam misalignment. Otherwise, MB-NOMA outperforms SB-NOMA and has a constant sum-rate performance, as obtained in Lemma 3.

Lemma 3 Using LSAAAs with a large number of M_x and M_z but finite or in a high SNR regime, MB-NOMA has the following constant sum-rate performance $R_{\text{asym}}^{\text{MB}}$ when $\Delta \vec{\Theta} > \vec{\Omega}_1^{\text{3dB}}$ for the

¹With 128×1 ULA, the 3dB-width Ω_1^{3dB} of the beam generated at $\theta_1 = 10^\circ$ is equal to $\Omega_1^{\text{3dB}} = 2.25^\circ$.

2-user scenario

$$R_T^{MB} \xrightarrow[M \gg 1]{M_x \gg 1, M_z \gg 1} \log_2 \left(1 + \frac{\gamma_{1,g} \rho |\alpha_{1,1}|^2 M}{4} \right) + \log_2 \left(1 + \frac{1 - \gamma_{1,g}}{\gamma_{1,g}} \right) = R_{asym}^{MB}. \quad (5.25)$$

Proof 3 Using LSAAs with a large number of M_x and M_z , $\mathbf{a}^H(\vec{\Theta}_{1,1}, M_x^{sa}, M_z^{sa}) \mathbf{a}(\vec{\Theta}_{1,2}, M_x^{sa}, M_z^{sa}) \xrightarrow[M_x \gg 1, M_z \gg 1]{M_x \gg 1, M_z \gg 1} 0$. Thereby, by setting $\beta'_{1,2}(M_x^{sa}, M_z^{sa}) \xrightarrow[M_x \gg 1, M_z \gg 1]{M_x \gg 1, M_z \gg 1} 0$, (5.23) can be rewritten by

$$R_T^{MB} = \log_2 \left(1 + \frac{\gamma_{1,g} \rho |\alpha_{1,1}|^2 M}{4} \right) + \log_2 \left(1 + \frac{\rho(1 - \gamma_{1,g}) |\alpha_{1,2}|^2 M}{\rho \gamma_{1,g} |\alpha_{1,2}|^2 M + 4} \right) \quad (5.26)$$

The second term in (5.26) can be approximated to $\log_2 \left(1 + \frac{1 - \gamma_{1,g}}{\gamma_{1,g}} \right)$ in a high SNR regime, i.e., $\rho \gg 1$ or with a large number of transmit antennas but finite, i.e., $M \gg 1$. Thus, Lemma 3 is verified.

The corresponding value of R_{asym}^{MB} in (5.25) is approximately equal to 17.6 and 19.4, using ULA with $M = 128$ and 16×16 URA, respectively. From Fig. 5.4a and Fig. 5.5, we find that the simulation results verify the expression in Lemma 3 for ULA and URA, respectively.

Remark 3 Based on the aforementioned analysis, we conclude that SB-NOMA is suitable only when users are very close to each other. Otherwise, MB-NOMA has a significant constant sum-rate performance independently of $\Delta\vec{\Theta}$, compared to SB-NOMA. Interestingly, from the obtained results and by extending to the more general multi-user scenario, it is preferable to cluster first the users with high spatial interference within SB-NOMA groups and then cluster the remaining users in MB-NOMA groups.

Remark 4 SDMA has a significant sum-rate performance with LSAAs, when the users are well separated in the angle domain. However, for HBF-based m -MIMO-NOMA systems, SDMA fails to accommodate all the users in overloaded scenarios, i.e., when $K > N_{RF}$. Indeed, SDMA can accommodate one user at each RF chain. Alternatively, MB-NOMA can serve multiple users with arbitrary AoDs within one RF chain and has a considerable system sum-rate.

5.4.3 System model of the multi-user joint SB- and MB-NOMA scheme

In this section, we propose a two-stage UC algorithm for the multi-user hybrid AD MIMO-NOMA systems. In the first stage, we group the users having high spatial interference within the same SB-NOMA group according to β -SB-NOMA UC in Section 4.4.2. It is worth noting that this UC only requires the AoD of users. In the second stage, we partition the remaining users, either 2-by-2 within MB-NOMA or alone in SU groups. For the MB-NOMA framework, and according to Subsection 5.4.2, users with arbitrary AoDs can be served within the same group, i.e., by the same RF chain. So, finally, in the second stage, MB-NOMA UC partitioned the G_{SU} users into G'_{SU} SU groups and G_{MB} 2-user MB-NOMA groups so that the total number of groups $G_f = G_{SB} + G'_{SU} + G_{MB}$ satisfies $G \leq N_{RF}$. In this thesis, we propose two schemes, namely SB-MB-NOMA and SB-MB-NOMA-RRFC, where the difference lies in the second stage.

Fig. 5.7 illustrates the system model of the proposed schemes and Fig. 5.8 plots the flowchart of each scheme.

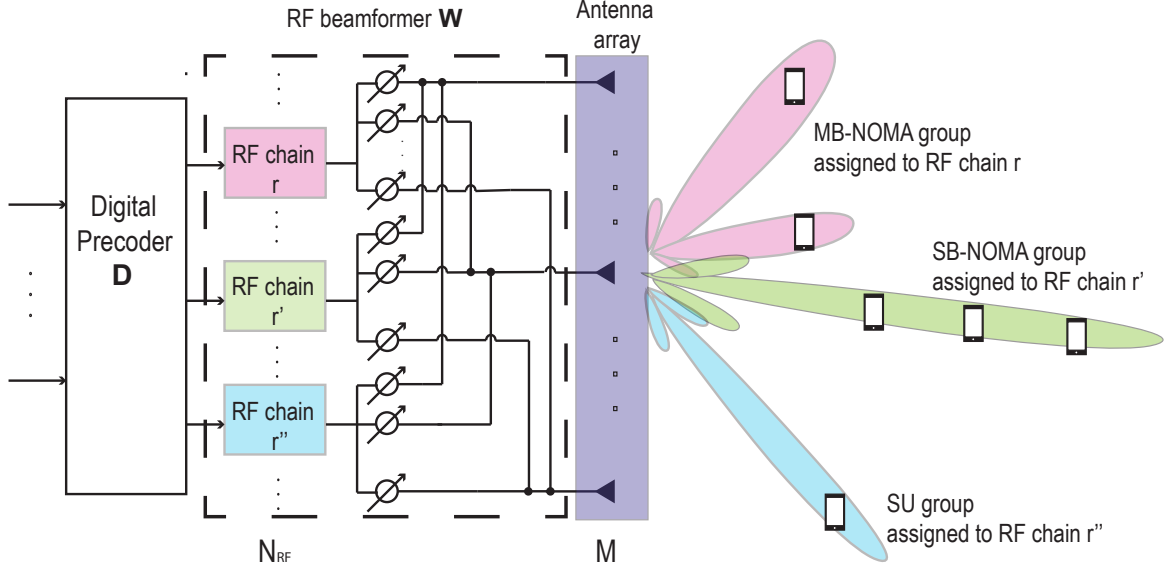


Figure 5.7: Joint SB- and MB MIMO-NOMA scheme.

The transmit RF beamformer $\mathbf{w}_g$ related to group g depends on its type and is given by

$$\mathbf{w}_g = \begin{cases} \mathbf{a}(\vec{\Theta}_g, M_x, M_z) \text{ with } \vec{\Theta}_g \text{ calculated as in (5.4)} & \text{if SB-NOMA with } \text{card}(\mathcal{S}_g) \geq 2, \\ \mathbf{w}_g \text{ is calculated as in (5.11) and (5.13)} & \text{if MB-NOMA with } \mathcal{S}_g = \{k, k'\}, \\ \mathbf{a}(\vec{\Theta}_{1,k}, M_x, M_z) & \text{if SU with } \mathcal{S}_g = \{k\}. \end{cases} \quad (5.27)$$

5.4.3.1 SB-MB-NOMA scheme

In the SB-MB-NOMA scheme, we aim to improve the SE by leveraging the potentiality of SDMA with LSAAs and the potentiality of SB-NOMA when the users are well close to each other. And, we apply MB-NOMA only in an overloaded scenario, i.e., if $G_{SU} + G_{SB} > N_{RF}$ after the first stage. Specifically, we partition the obtained G_{SU} users within G_{MB} MB-NOMA groups and G'_{SU} SU groups, i.e., $G_{SU} = 2G_{MB} + G'_{SU}$. In the second stage, we group 2-by-2 the G_{SU} users in 2-user MB-NOMA groups until all the obtained G_f groups can be served simultaneously by the BS. In other words, $G_f = G_{SB} + G_{MB} + G'_{SU} = N_{RF}$. Accordingly, the number of MB-NOMA groups that must be generated is given by $G_{MB} = G_{SU} + G_{SB} - N_{RF} > 0$.

5.4.3.2 SB-MB-NOMA-RRFC scheme

In the SB-MB-NOMA-RRFC scheme, we aim to **R**educe the number of the active **RF** **C**hains as much as possible. To do so, we group all the G_{SU} users 2-by-2 in MB-NOMA groups. Hence, $G_{MB} = \lfloor \frac{G_{SU}}{2} \rfloor$ and $G'_{SU} = 0$ (resp. 1) if G_{SU} is even (resp. odd). The optimal combination of MB-NOMA groups that maximizes the sum-rate can be obtained by exhaustive search over all

the possible combinations, whose number is equal to $\prod_{m=1}^{C_{SU}/2} (2*m-1)$ (resp. $\prod_{m=1}^{(C_{SU}+1)/2} (2*m-1)$) if C_{SU} is even (resp. odd).

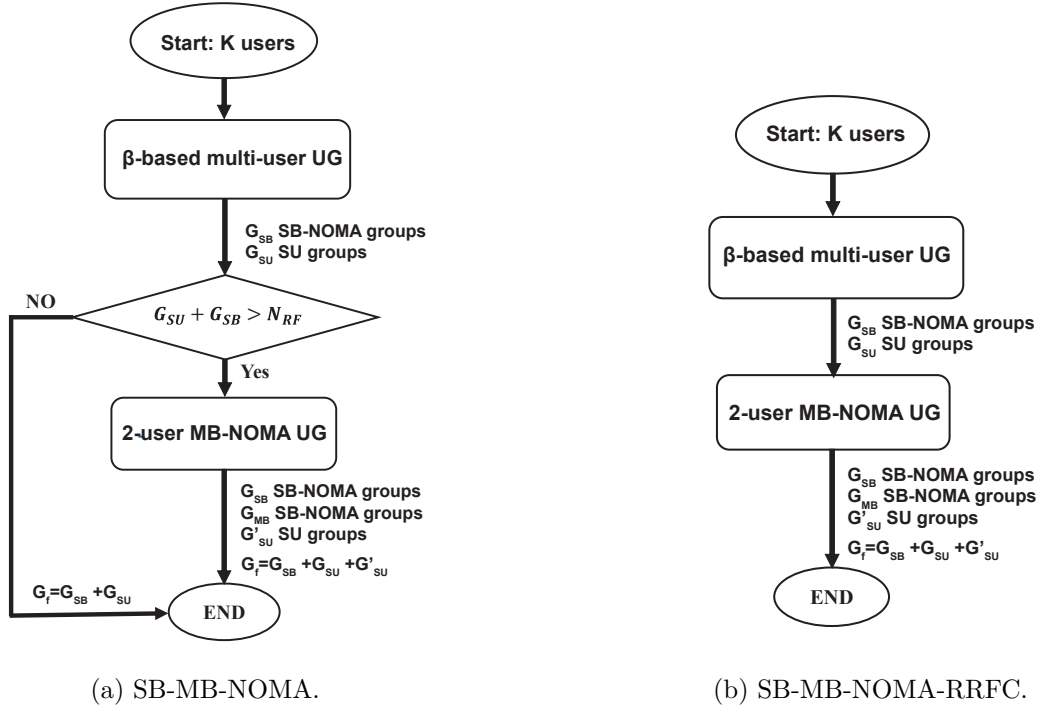


Figure 5.8: Flowchart of the proposed SB-MB-NOMA and SB-MB-NOMA-RRFC schemes.

5.4.4 Low-complex MB-NOMA group clustering

To reduce the complexity of both schemes, we develop three MB-NOMA UC strategies. From Subsection 5.4.2, once the users having high spatial interference are clustered into SB-NOMA groups, the choice of associating two users in MB-NOMA according to their AoD is not critical regarding the sum-rate. To this end, we first sort the G_{SU} users in an array in the ascending order of their azimuth AoD, i.e., θ for both ULA and URA. Then, we propose three different selection strategies as shown in Fig. 5.9, denoted as UC1, UC2, and UC3, to select 2-by-2 the users from the sorted array and cluster them in MB-NOMA groups until all the G_{MB} MB-NOMA groups are given. Recall that the value of G_{MB} is calculated according to the scheme type.

For SB-MB-NOMA-RRFC scheme using a ULA with $N_{BS=128}$ antennas, the sum-rate performance of the proposed MB-NOMA UC strategies is compared in Fig. 5.10 with the optimal exhaustive search method, denoted as OPT UC, and the random MB-NOMA UC, where the users are randomly clustered in MB-NOMA groups. Because of memory size limit, we plot the sum-rate of the OPT UC until $K = 10$. Note that the particular form of the curves, especially for a low number of users, is not a result of the number of realizations but related to the number of users being odd or even. Indeed, for a low number of users, it is more probably that users are well separated in the angle domain and then all are clustered in SU groups after the first stage, i.e., $G_{SU} = K$. Thus, if K is even (resp. odd), there exist $G'_{SU} = 0$ (resp. 1) user served via

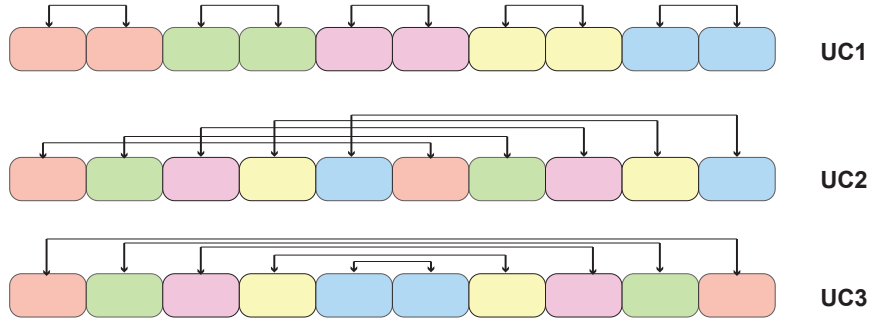


Figure 5.9: Three different 2-by-2 selection strategies.

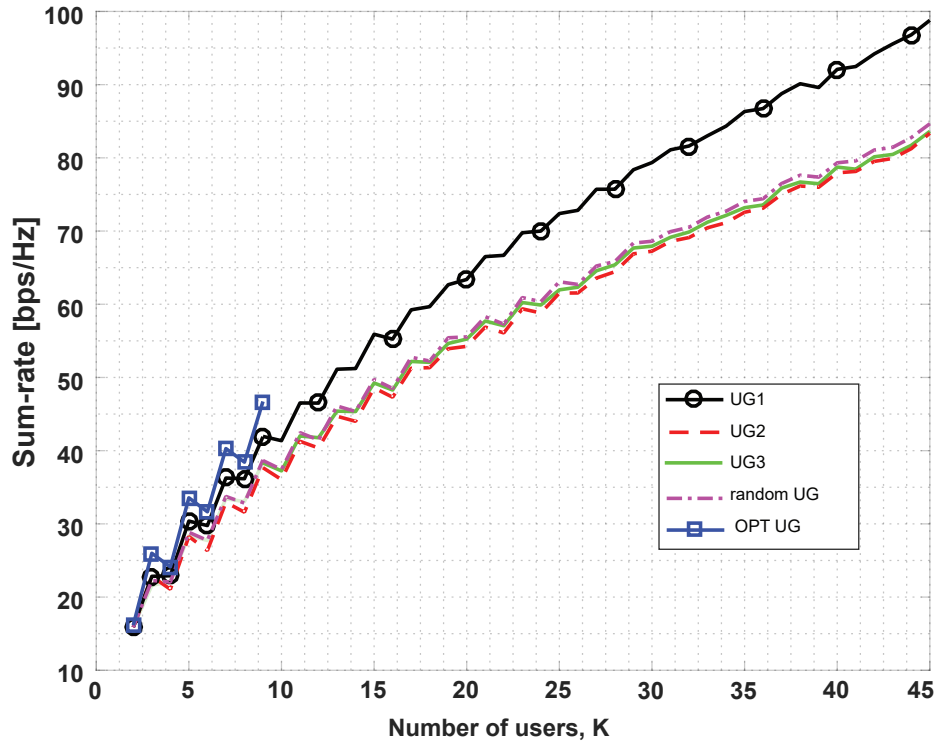


Figure 5.10: The sum-rate performance of different MB-NOMA UC algorithms versus the number of users for the SB-MB-NOMA-RRFC scheme. The sum-rate is averaged over 5,000 trials.

SDMA. For that, it is obvious that the sum-rate decreases slowly when passing from $K = 2 * Q$ to $K = 2 * Q + 1$ in case of sparse cell, because of the potential of SDMA with LSAs which enables high beamforming gain. Moreover, it is seen from Fig. 5.10 that UC1 is closer to OPT UC than the others. This superiority is because of the lower inter-cluster interference achieved when using UC1 compared to UC2 and UC3. In Section 5.5, UC1 is considered as the reference.

5.5 Illustrative results and discussions

In this section, we evaluate the performance of the proposed schemes for mmWave hybrid systems via simulation. Specifically, we consider a mmWave rural environment using NYUSIM. The

specific values of the simulation parameters adopted in this chapter are summarized in Table 5.1.

Table 5.1: Simulation parameters

Parameters	Value
Number of transmit antennas, M	$\in \{128, 256\}$
Number of RF chains, N_{RF}	$\in \{20, 25\}$
Carrier frequency	28 GHz
Channel bandwidth	20 MHz
Cell radius	100 m
Total transmission power, P_e	30 dBm
Noise power, σ_n^2	-101 dBm
Minimum SIC power difference, $P_{\min}$	1 mW
Number of paths per time cluster in a rural environment	$\in \{1, 2\}$

To illustrate the effectiveness of our proposed schemes, we adopt three baseline schemes; two apply fully DBF (thus requiring $N_{RF} = M$) and the other uses HBF with $N_{RF} \leq M$. The following acronyms will be used to refer to the different schemes

- AD-DBF: denotes the scheme considered in Chapter 3, where the BS generates K directive beams, each one is steered in the AoD of the intended user.
- AD-DBF-SB-NOMA: denotes the scheme proposed in Chapter 4, where the β -SB-NOMA UC algorithm designed in Section 4.4.2 clusters the users within SU and multi-user SB-NOMA groups.
- AD-HBF-SB-NOMA-TDMA: denotes the scheme proposed in [25], where the authors propose a simple AD UC strategy, in which UEs with the same estimated angle belong to the same group². However, due to the limited RF chains, the authors developed a group clustering algorithm in the time domain that reduces the inter-group interference by clustering the groups with small angular distances in distinct time slots.
- β -HBF-SB-NOMA-TDMA: denotes the scheme proposed in Section 5.3.
- AD-HBF-SB-MB-NOMA: denotes the scheme proposed in Subsection 5.4.3.1.
- AD-HBF-SB-MB-NOMA-RRFC: denotes the scheme proposed in Subsection 5.4.3.2.

For a fair comparison, the different schemes apply the same user ordering and power allocation strategies proposed in Chapter 4 that require only angular information, i.e., users' AoD. Accordingly, each scheme only requires the AoD of each user.

5.5.1 2D HBF-based MIMO-NOMA performance

Fig. 5.11 illustrates the average sum-rate and the average number of active RF chains, N_{RF}^a , per time slot versus the number of users for the different aforementioned schemes. The BS adopts

²Note that the estimated angles belong to a predefined azimuth angle set with a fixed search step size $J = M$ [25].

a ULA array with $M = 128$ transmit antennas to serve the $2 \leq K \leq 2N_{RF}$ users. For HBF, we assume $N_{RF} = 20$ RF chains at the BS. In contrast, the fully DBF involves $M = 128$ RF chains.

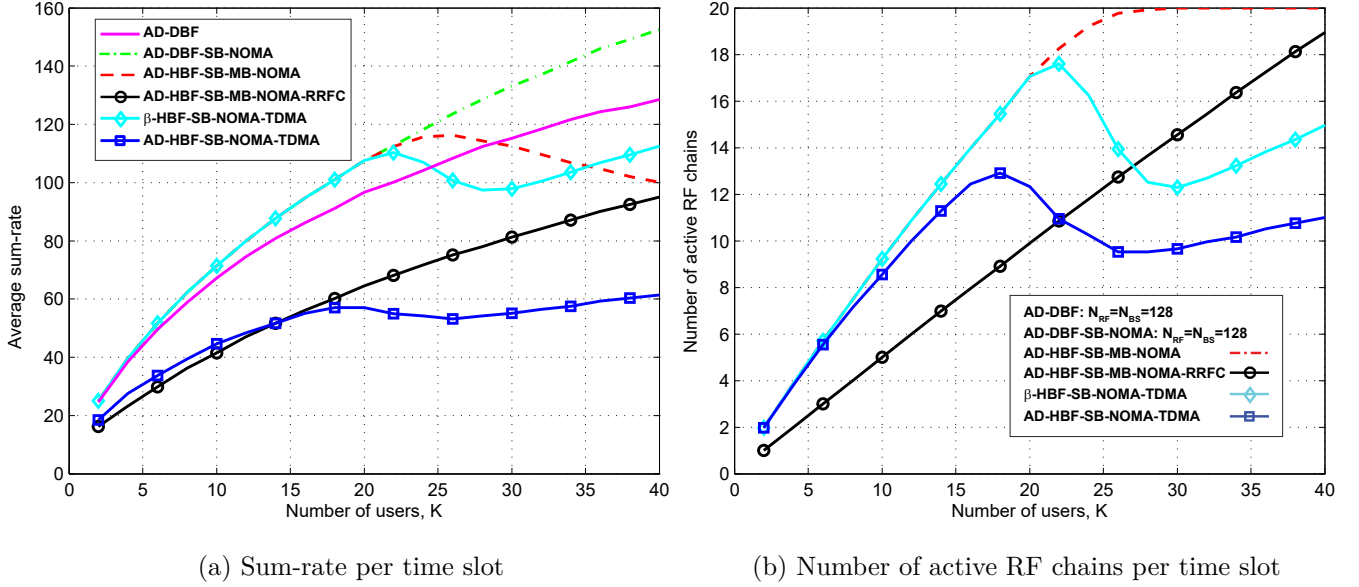


Figure 5.11: (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users with a 128×1 ULA array and 20 RF chains at the BS.

For $K \leq N_{RF}$, AD-HBF-SB-NOMA (represented by AD-HBF-SB-MB-NOMA and HBF-SB-NOMA-TDMA) is preferable to AD-DBF-SB-NOMA since the hybrid scheme has the same performance as the digital one with lower complexity, cost, and energy consumption. However, their corresponding curves separate for $N_{RF} < K \leq 2N_{RF}$. Indeed, when $K \leq N_{RF}$, the users are well separated in the space with a high BF gain brought by the narrow beam of LSAA's. Thereby, AD-HBF-SB-NOMA doesn't exploit the potentiality of NOMA to add more DoFs when $K > N_{RF}$. Interestingly, the TDMA and the MB-NOMA frameworks, respectively, in the proposed AD-HBF-SB-MB-NOMA and β -HBF-SB-NOMA-TDMA schemes, offer new DoFs to compensate for the few RF chains and provide user connectivity even if $K > N_{RF}$. Besides, as seen in Fig. 5.11, AD-HBF-SB-MB-NOMA and β -HBF-SB-NOMA-TDMA provide a good performance compared to the AD-DBF scheme with $128/20 \leq M/N_{RF} \leq 128/2$ times fewer active RF chains. More details of the evolution of the sum-rate for each scheme will be analyzed separately in the next paragraphs.

Performance of β -HBF-SB-NOMA-TDMA

It can be seen from Fig. 5.11a that the sum-rate of β -HBF-SB-NOMA-TDMA for $N_{RF} < K \leq 2N_{RF}$ first decreases, then increases as K becomes larger. Indeed, for $N_{RF} < K \leq 2N_{RF}$, TDMA is applied in the second phase to serve the SU and SB-NOMA groups obtained in the first phase at different time slots. Serving the users in different time slots reduces the system sum-rate and the necessity for more active RF chains. But then adding more users improves the sum-rate. The more the users are served simultaneously, the more the sum-rate is improved. For that, it

can be seen that the evolution of the sum-rate per time-slot matches with that of the number of active RF chains per time-slot in Fig. 5.11b.

Furthermore, Fig. 5.11a shows that β -HBF-SB-NOMA-TDMA scheme outperforms the one proposed in [25]. For example, when $K = 2N_{RF} = 40$ users, β -HBF-SB-NOMA-TDMA achieves a sum-rate gain up to 83% over AD-HBF-SB-NOMA-TDMA. Indeed, the latter [25] uses a predefined angle set and the users having the same estimated angle belong to the same SB-NOMA groups. Subsequently, the authors cluster the obtained groups in the time domain according to their angular distance. The results in Fig. 5.11a reveal the performance of our proposed user and group clustering algorithms in SB-NOMA groups and the time domain, respectively, against those proposed in [25]. In addition, HBF-SB-NOMA-TDMA is proposed for ULA, which is not feasible in practice. Interestingly, our proposed β -HBF-SB-NOMA-TDMA scheme uses the spatial interference metric β built based on the array factor. Thus, β -HBF-SB-NOMA-TDMA can be applied for any antenna array architecture, which can be ULA, UCA (uniform circular array), URA, etc. The performance evaluation with URA will be illustrated in the next section.

Performance of AD-HBF-SB-MB-NOMA

The use of TDMA with NOMA imposes more signal processing and synchronization complexity at the BS and the receiver sides. To address this issue, another scheme, namely AD-HBF-SB-MB-NOMA, is proposed to offer more DoFs using only NOMA. Importantly, this scheme outperforms AD-DBF for $K < 28$ with $128/20 \leq M/N_{RF} \leq 128/2$ times fewer RF chains. Otherwise, AD-DBF (with or without NOMA) has a higher sum-rate than AD-HBF-SB-MB-NOMA at the expense of high cost and prohibitive complexity. Indeed, while the number of active RF chains, N_{RF}^a , for AD-DBF (with or without NOMA) is always equal to the number of antennas, $M = 128$, it will be smaller than or equal to $N_{RF} = 20$ for AD-HBF-SB-MB-NOMA, as seen in Fig. 5.11b.

Fig. 5.11a shows a monotonical increase of the sum-rate of AD-HBF-SB-MB-NOMA with the number of users until $K = 26$, then a performance decrease. This evolution can be explained by Fig. 5.11b, which illustrates that the number of active RF chains, N_{RF}^a , increases with the number of users until $K = 26$, then saturates at $N_{RF}^a = N_{RF}$. In other words, for $26 < K \leq 40$, the number of MB-NOMA groups and that of SU groups increases and decreases, respectively, as K increases, to satisfy that $G_f = N_{RF}$. It is worth noting that the BF gain at each user served within the MB-NOMA group is halved compared to that at the user served using a single beam within the SU group, as shown in Fig. 5.3. For that, for $26 < K \leq 40$, the performance gap between AD-DBF-SB-NOMA where each antenna is attached to a RF chain and AD-HBF-SB-MB-NOMA with few RF chains (relative to the antennas number) increases with the number of users, K . Note that AD-HBF-SB-NOMA with N_{RF} RF chains can not simultaneously support all the users if $K \geq N_{RF}$, while AD-HBF-SB-MB-NOMA allows for their communication with a considerable performance. Interestingly, AD-HBF-SB-MB-NOMA, with a basic antenna allocation (divides the entire array into two equal subarrays) and a low-complex MB-NOMA

user clustering strategy, achieves good performance with very few active RF chains compared to the digital schemes. To further enhance the system sum-rate, one good perspective is to design efficient antenna allocation and user clustering techniques that optimize the BF gain at each user according to their position against the others and reduce the inter-group interference.

Moreover, AD-HBF-SB-MB-NOMA has a significant performance gain over the previous work [25], by using only NOMA. For instance, when $K = 25$ (resp. $K = 40$), AD-HBF-SB-MB-NOMA outperforms AD-HBF-SB-NOMA-TDMA by almost 99% (resp. 66.6%). In addition, AD-HBF-SB-MB-NOMA outperforms β -HBF-SB-NOMA-TDMA scheme for $N_{RF} < K < 35$. Then, for $35 < K < 40$, β -HBF-SB-NOMA-TDMA slightly outperforms with less active RF chains at the expense of high signal processing and synchronization at the BS and the receiver sides. For example, β -HBF-SB-NOMA-TDMA achieves a sum-rate gain of 10% over AD-HBF-SB-MB-NOMA when $K = 40$.

Performance of AD-HBF-SB-MB-NOMA-RRFC

Fig. 5.11a shows that when $K \leq N_{RF}$, the sum-rate gap between AD-HBF-SB-MB-NOMA and AD-HBF-SB-MB-NOMA-RRFC increases with the number of users, K . Indeed, the latter purposes reduce the number of active RF chains and cluster all the remaining users in 2-user MB-NOMA groups. In contrast, the former aims to improve spectral efficiency and serve the remaining users, each with a single beam having a high BF gain, when $K \leq N_{RF}$. These can be illustrated by the number of active RF chains in Fig. 5.11b. Otherwise, for $N_{RF} < K \leq 2N_{RF}$, this gap decreases as K increases, and vanishes approximately at $K = 2N_{RF}$. Indeed, it is obvious that for $N_{RF} < K \leq 2N_{RF}$, more users are clustered in MB-NOMA groups as K increases so that the total number G_f of groups satisfies that $G_f \leq N_{RF}$.

Besides, it can be seen that AD-HBF-SB-MB-NOMA-RRFC reduces the number of active RF chains more than β -HBF-SB-NOMA-TDMA and AD-HBF-SB-NOMA-TDMA for $K \leq 27$ and $K \leq 22$, respectively. Otherwise, both TDMA schemes have lower N_{RF}^a at the expense of additional signal processing complexity at the BS and the receiver sides. This limitation is because only two users can be served in MB-NOMA groups. To more reduce the energy consumption, the size of MB-NOMA group can exceed two, and this consideration will be left as a future work.

As a conclusion, AD-HBF-SB-MB-NOMA improves the spectral efficiency with a limited number of RF chains without applying additional MA technique, and AD-HBF-SB-MB-NOMA-RRFC succeeds in reducing the number of active RF chains, i.e., the energy consumption, with only two users in each MB-NOMA group.

5.5.2 3D HBF-based MIMO-NOMA performance

URA is more practical for massive MIMO than ULA [150]. Indeed, unlike ULA, URA can pack a massive number of antennas within a small 2D area. On the other hand, URA can estimate both the azimuth and the elevation angles, unlike ULA that finds only one information. To this end, in this section, we study the case when the BS adopts URA. Specifically, we investigate the

impact of different factors on the proposed schemes, such as the antenna array architecture and the number of transmit antennas and RF chains.

5.5.2.1 Impact of the antenna array architecture

URA with 128 antennas at the BS

Here, we highlight the impact of the antenna array architecture. First, we assume that M and N_{RF} are equal to that considered in Section 5.5.1 using ULA, i.e., $M = 128$ and $N_{RF} = 20$, but with different antenna array architectures, namely 16×8 URA and 32×4 URA. And, we plot in Fig. 5.12 the average sum-rate and the average number of active RF chains versus the number of users, K .

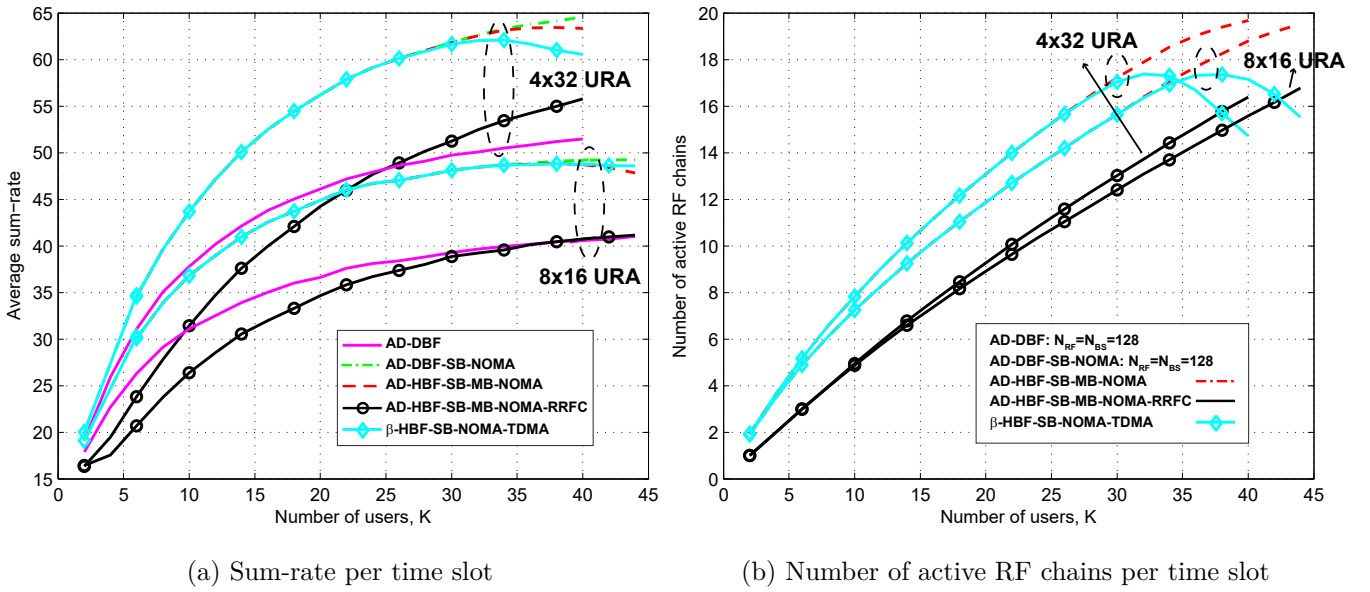


Figure 5.12: (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users of the different schemes with 16×8 URA and 32×4 URA and 20 RF chains at the BS.

It's clear from Fig. 5.11 and Fig. 5.12 that the point I_d , at which the curves of AD-DBF-SB-NOMA, AD-HBF-SB-MB-NOMA, and β -AD-HBF-SB-NOMA-TDMA differentiate, depends on the array architecture. For instance, it is at $K = N_{RF} = 20$, $K = 30$, and $K = 38$ for 128×1 ULA, 32×4 URA, and 16×8 URA, respectively. Indeed, when more antennas lie in the azimuth plane, the beam becomes narrower, and thereby SB-NOMA serves fewer users. Note that with URA, the 3D beam is wider than the 2D beam with ULA using the same number of antennas, $M = 128$, and for that, with URA, SB-NOMA can offer additional DoFs for some $K > N_{RF}$.

As can be observed from Fig. 5.12, using URA, AD-HBF-SB-MB-NOMA and β -HBF-SB-NOMA-TDMA with few active RF chains outperform AD-DBF where each antenna is attached to one RF chain. For instance, using 32×4 URA (resp. 16×8 URA), AD-HBF-SB-MB-NOMA outperforms AD-DBF by almost 23% (resp. 16.5%), when $K = 40$ (resp. $K = 44$) users are served in the cell with $128/20 = 6.4$ times fewer active RF chains. In addition, AD-HBF-SB-

MB-NOMA-RRFC has a similar sum-rate as AD-DBF for $30 \leq K \leq 44$ and $25 \leq K \leq 40$ using 16×8 URA, and 32×4 URA, respectively, with few active RF chains as shown in Fig. 5.12b. Moreover, since only two users can be served in MB-NOMA groups, the MB-NOMA schemes can not exploit the multi-user diversity for $K > 40$ and $K > 44$, using respectively 32×4 URA and 16×8 URA.

The results reveal that our proposed schemes have an even more significant interest with URA than with ULA because the beamwidth is wider with URA. This includes high spectral efficiency and lower number of active RF chains, i.e., lower energy consumption, compared to the DBF-based schemes, namely AD-DBF and AD-DBF-SB-NOMA.

URA with 256 antennas at the BS

Now, we will exploit the possibility of implementing a massive number $M = 256$ of antennas at the BS using URA, and we adopt a 32×8 URA at the BS. For HBF, we consider $N_{RF} = 20$ RF chains at the BS. Fig. 5.13 plots the average sum-rate and the average number of active RF chains for the different schemes versus the number of users, K .

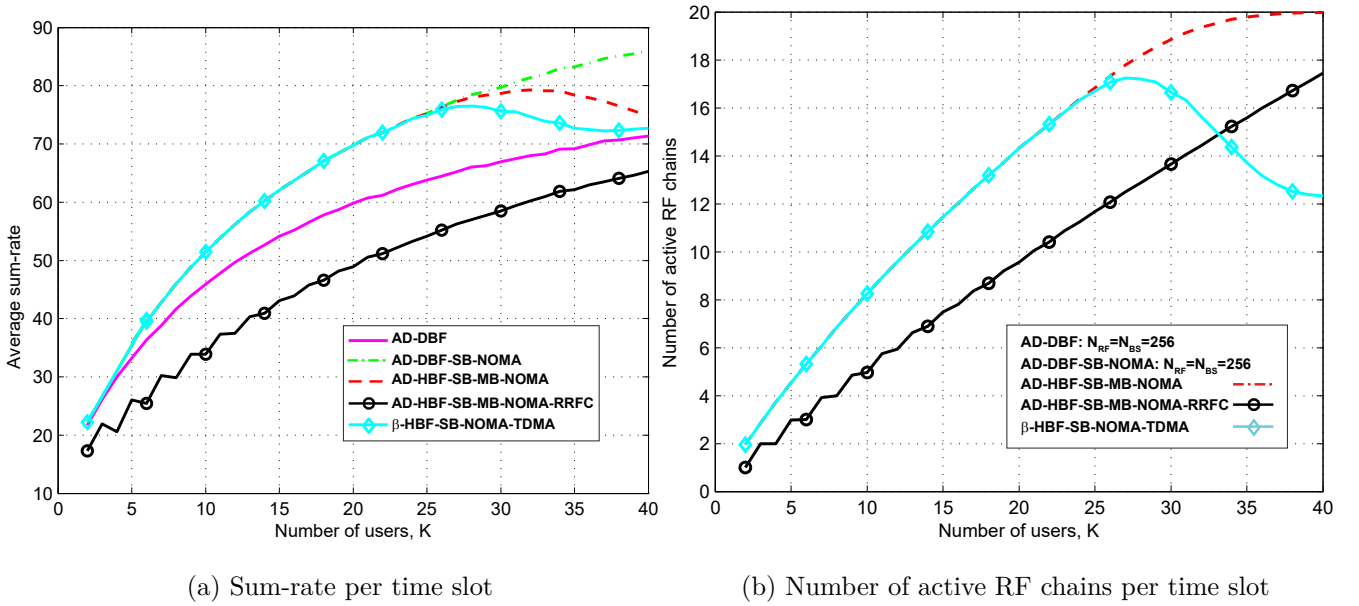


Figure 5.13: (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of users with 32×8 URA and 20 RF chains at the BS.

Increasing the antennas number decreases the beamwidth, and thereby fewer users are within SB-NOMA groups. Fig. 5.13 illustrates this fact from point I_d that stands at $K = 25$ for 32×8 URA. Note that with URA and even with a twice number of antennas of that with ULA, there are more users clustered in SB-NOMA groups. The reason is that the 3D beamwidth depends on the number of both azimuth and the elevation antennas and not only on the total number of antennas. Interestingly, we can observe from Fig. 5.13 that the three proposed schemes have an interesting performance gain against AD-DBF and AD-DBF-SB-NOMA for $2 < K \leq 2N_{RF}$ using URA when compared to ULA in Fig. 5.11. Simulation results show that the proposed

scheme obtains a much lower hardware complexity (e.g., $N_{RF} = 20 \ll M = 256$) with only a slight performance degradation compared to AD-DBF-SB-NOMA, using URA. For instance, for $K = 2N_{RF} = 40$, β -HBF-SB-NOMA-TDMA and AD-HBF-SB-MB-NOMA degrade with a 13.6% decrease gain w.r.t. AD-DBF-SB-NOMA with respectively 256/12 and 256/20 fewer active RF chains.

5.5.2.2 Impact of the antennas number

Varying the number of horizontal antennas

We fix the number of vertical antennas $M_z = 8$ and we vary the number of horizontal antennas, M_x . And, we plot in Fig. 5.14 the average sum-rate and the average number of active RF chains when the BS has $N_{RF} \in \{20, 25\}$ RF chains and serves $K = 40$ users.

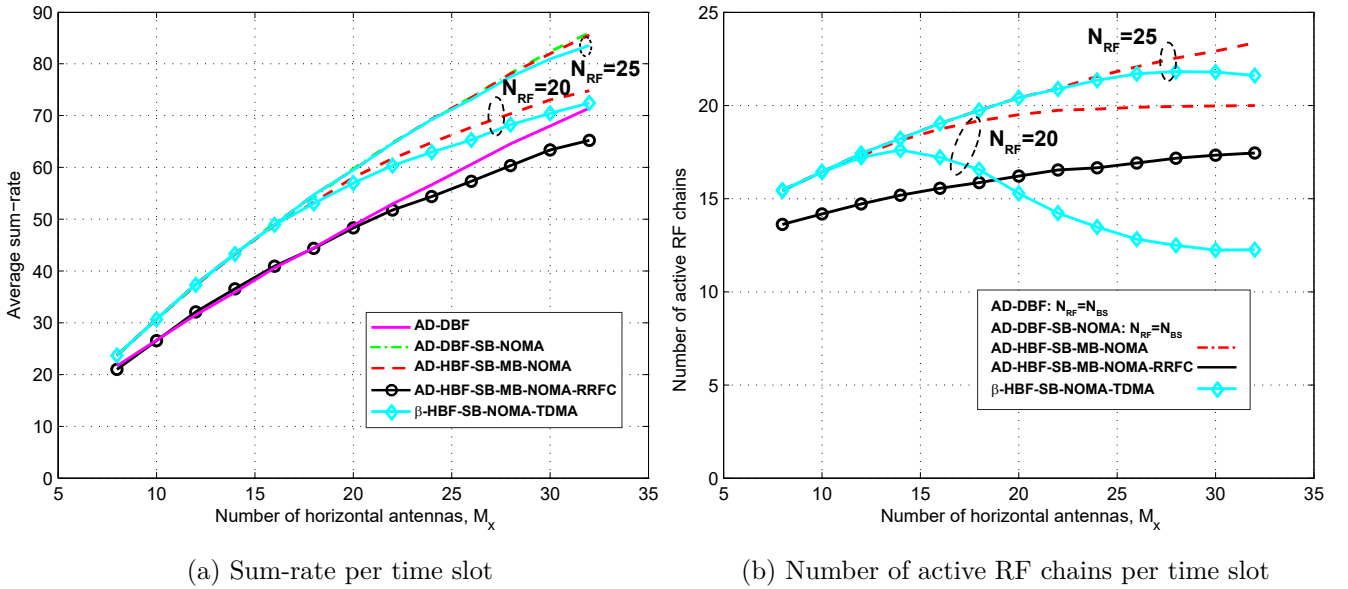


Figure 5.14: (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of horizontal antennas, M_x , when $K = 40$ users exist in the cell with $M_z = 8$ vertical antennas and $N_{RF} \in \{20, 25\}$ RF chains at the BS.

It is observed in Fig. 5.15 that the sum-rate increases as the number of horizontal antennas, M_x , becomes bigger. An intuitive explanation is that implementing more transmit antennas can provide higher BF gain and more spatial diversity gains, thus enhancing spectral efficiency. For instance, with $N_{RF} = 20$, the $K = 40$ users are served using only SB-NOMA when $8 \leq M_x \leq 14$ horizontal antennas exist at the BS due to a sufficient large beamwidth. Instead, increasing the horizontal antennas enables more accurate user discrimination in azimuth, and few users belong to SB-NOMA groups. Accordingly, the curves of AD-DBF-SB-NOMA, AD-HBF-SB-MB-NOMA, and β -AD-HBF-SB-NOMA-TDMA separate. And, the sum-rate gap between these schemes increases with M_x .

Fig. 5.15 shows that raising the number of RF chains, N_{RF} , can considerably improve the sum-rate performance. Indeed, the BS can simultaneously serve more users with more RF

chains. Note that AD-HBF-SB-MB-NOMA-RRFC has the same performance whatever N_{RF} since $K = 40 \leq 2N_{RF} \forall N_{RF} \in \{20, 25\}$.

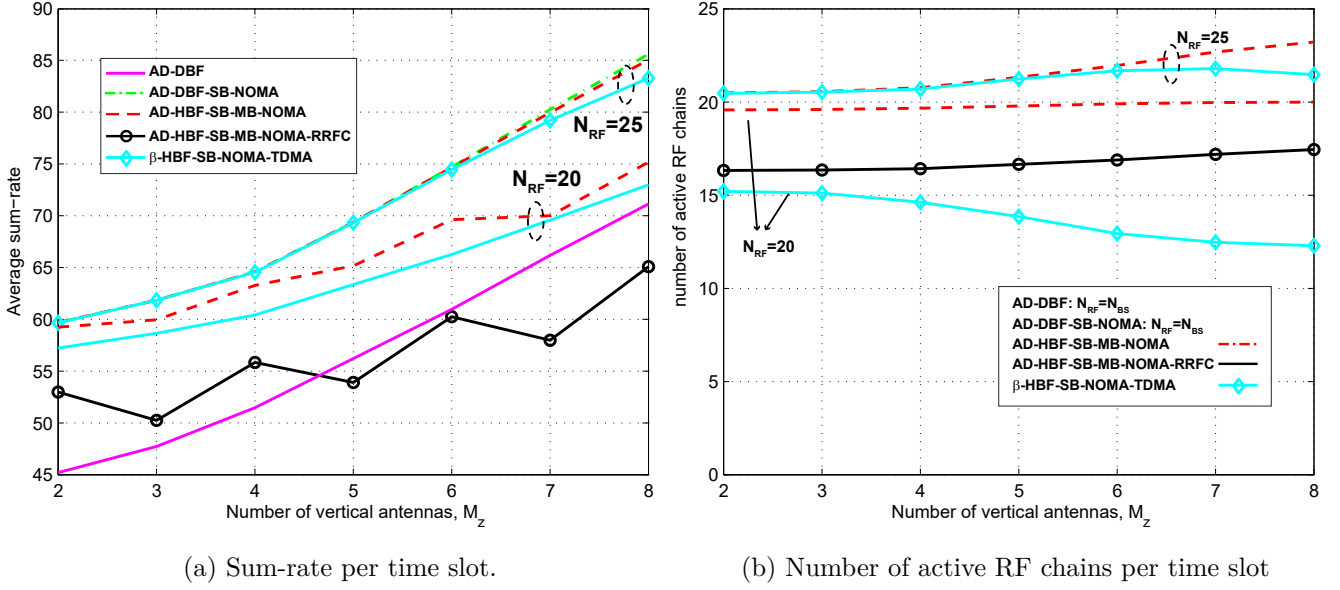


Figure 5.15: (a) The average sum-rate and (b) the average number of active RF chains per time slot versus the number of vertical antennas M_z when $K = 40$ users exist in the cell with $M_x = 32$ horizontal antennas and $N_{RF} \in \{20, 25\}$ RF chains at the BS.

Varying the number of vertical antennas

Now, we fix the number of horizontal antennas $M_x = 32$ and we vary the number of vertical antennas M_z . And, we plot in Fig. 5.15 the sum-rate and the number of RF chains for $N_{RF} \in \{20, 25\}$ when $K = 40$ users exist in the cell.

Since $M_x = 32$ horizontal antennas, the BS highly discriminates the users in azimuth. For that, we find that, with $N_{RF} = 20$, SB-NOMA can not support the $K = 40$ users even with only $M_z = 2$ vertical antennas. Moreover, the achievable sum-rate of the different schemes still increases as M_z evolves greater due to the high BF gain. In addition, the gap between the proposed schemes and AD-DBF-SB-NOMA becomes wider due to the narrow beam. Similarly to Fig. 5.14, we obtain that the HBF-based MIMO-NOMA schemes have a significant sum-rate performance with more RF chains and that AD-HBF-SB-MB-NOMA-RRFC has the same performance for $N_{RF} \in \{20, 25\}$ since $K = 40 \leq 2N_{RF}$.

5.6 Summary

The main objective of the current chapter is to overcome the restriction incurred by the few RF chains in mmWave hybrid systems and propose HBF-based mmWave m-MIMO-NOMA schemes with only the knowledge of the angular information. This chapter focuses on mmWave HBF-based m-MIMO-NOMA systems, where only angular information is estimated. Specifically, we

address the limitation of SB-NOMA schemes to exploit the multi-user diversity in mmWave hybrid systems with LSAAs. And, we have proposed two different two-phase schemes to offer additional DoFs. The former adopts TDMA and the latter leverages the potential of MB-NOMA. Unlike the prior works, the proposed schemes can be applied to both ULA and URA architectures, and require only angular information. Simulation results have shown that the proposed schemes yield significant performance gains in terms of sum-rate and number of active RF chains, when compared to the solution proposed in the literature and other schemes based on fully DBF. For instance, the proposed TDMA-based scheme achieves a sum-rate gain of up to 83% over the TDMA-based one existing in the literature. Moreover, using 32×4 URA, AD-HBF-SB-MB-NOMA outperforms AD-DBF by almost 23%, when $K = 40$ users are served in the cell with 6.4 times fewer active RF chains.

Chapter 6

Conclusion and Perspectives

This chapter summarizes the contributions of this dissertation and points out future research directions.

6.1 Thesis Summary

This dissertation addressed the critical challenges in mmWave m-MIMO-NOMA systems, namely, precoding design complexity, hardware constraints, channel estimation complexity and overhead, which result from using LSAAs. First, we exploited the sparse nature of the mmWave channels thanks to the high directionality and the potential blockage, which make the angular information a promising partial CSI. We then proved the interest of using only this angular information relative to the users' positions to perform low-complex precoding, user clustering and ordering, and power allocation techniques for mmWave m-MIMO-NOMA systems using both DBF and HBF techniques. To do so, we applied both 2D and 3D mmWave channels using the realistic and statistical mmWave channel NYUSIM simulator. Next, we present a summary of the research work that was carried out in this dissertation.

In Chapter 3, we studied the spatial behavior and potentiality of the AD-DBF technique in mmWave m-MIMO systems, using NYUSIM. As a reference, we considered CB that maximizes the BF gain at the intended user using a full CSI. We found that AD-DBF is a promising BF technique that balances performance, hardware complexity, and channel estimation overhead. In mono-path environments, AD-DBF gives a moderate rate for each user, while CB strengthens the strong users and marginalizes the weak ones. In multi-path environments, for AD-DBF, even though the users are well separated in space, there is interference resulting from NLoS paths. However, CB maximizes the BF gain by exploiting the diversity of NLoS paths, whereas AD-DBF lacks their potentiality and loses some of the radiated energy. With a large number of antennas relative to that of users, AD-DBF outperforms CB even in multi-path environments. Moreover, AD-DBF has a significant sum-rate performance in highly LoS environments, where all the paths depart from close directions.

In Chapter 4, we studied DBF-based MIMO-NOMA with dozens of antennas at the BS (e.g., 32, 64). And, we proposed a multi-user MIMO-NOMA transmission scheme, denoted as

AD MIMO-NOMA, at mmWave frequencies using only the angular information to address the estimation and feedback problems of FCSI. We first leverage the main features of mmWave channels, i.e., the high directionality and the potential blockage. Then, we derived a set of angular-based performance metrics, such as the inter-user spatial interference, the user channel quality, and the sum-throughput. These metrics then participated in the design of user clustering and ordering and power allocation techniques. Specifically, the user clustering algorithm groups the users with high spatial interference to perform NOMA. A new user ordering strategy is proposed using the angular-based channel quality metric. In addition, we designed a power allocation method that maximizes the angular-based sum-throughput. In this chapter, we studied not only the case of two-user NOMA groups, but also the case of multi-user NOMA groups. Extensive numerical results showed that the proposed multi-user AD MIMO-NOMA scheme significantly improves the performance of the mmWave MU-MIMO system by achieving up to 39% increase in the spectral efficiency when the number of users is close to that of antennas. Besides, we found that the proposed user ordering strategy outperforms other limited feedback strategies, and the angular-based power allocation allows for efficient SIC. And, the results showed the potential of NOMA against OMA to exploit the multi-user diversity in a congested cell. In addition, AD MIMO-NOMA can be an alternative solution to using a large antennas number. For instance, using URA, multi-user AD MIMO-NOMA with 64 antennas has the same performance as C-AD-DBF with 192 antennas.

In Chapter 5, we studied HBF-based m-MIMO-NOMA with hundreds of antennas at the BS (e.g., 128, 256). With LSAAs, the conventional SB-NOMA scheme cannot exploit the multi-user diversity in mmWave hybrid systems due to the narrowness of the beams and the restriction incurred by the few RF chains. In this chapter, we gave a particular attention to addressing these limitations, and we proposed two different schemes to offer additional DoFs for the conventional SB-NOMA scheme in mmWave hybrid systems. Interestingly, these schemes require only the knowledge of the angular information and apply with both ULA and URA architectures, contrary to those proposed in the literature. The former applies TDMA with SB-NOMA and exploits the time-domain resources by serving the clusters with high spatial interference within distinct time-slots. To reduce the complexity of signal processing at both Tx and Rx when applying TDMA and NOMA, we leveraged the MB-NOMA concept and proposed joint SB- and MB-NOMA schemes to benefit from NOMA multi-user diversity, whatever the load of the cell and the users' position are. Specifically, we proposed two joint SB- and MB-NOMA schemes: the former gives a significant sum-rate performance related to DBF, while the latter reduces the number of active RF chains, i.e., power consumption. We further validated the performance of the proposed schemes using numerical simulations in terms of sum-rate and the number of active RF chains by comparing them to the solution proposed in the literature and others using a fully DBF.

6.2 Future Research Perspectives

The proposed solutions in this dissertation addressed some of the challenges that arise in the context of precoding design complexity, hardware constraints, channel estimation complexity and overhead in next-generation cellular networks. Next, we present some open problems that can be addressed in the short term as an extension to both DBF- and HBF-based MIMO-NOMA schemes proposed throughout this dissertation

6.2.1 Short-term studies:

- To further improve the performance of the AD MIMO-NOMA schemes proposed in Chapter 4, it would be interesting to develop an iterative sub-optimal inter-cluster power allocation with the sub-optimal intra-cluster power allocation technique developed in Chapter 4. In addition, it is possible to design a power allocation strategy based on the angular information, that yields to a better fairness between users in the same cluster. Although, it would be interesting to consider the complexity and the imperfection of SIC decoding in the multi-user AD MIMO-NOMA schemes.
- All the schemes proposed throughout this dissertation assumed perfect channel estimation and feedback. However, this perfection is difficult to achieve in practice. Hence, there is a need to develop an error model of the angular information and to study the impact of imperfect angular information estimation and feedback on the performance of the proposed schemes. On the other hand, we consider continuous angles, which is not also feasible in practice especially with the analog phase shifters. Therefore, it would be interesting to study the proposed HBF-based schemes with low or high resolution analog phase shifters.
- The performance of the proposed schemes is studied in LoS and rural environments. Furthermore, it would be interesting to evaluate the performance in urban environments and with either LoS or NLoS scenarios. In the latter case, the LoS path w.r.t. the BS is blocked.
- In our work, we consider single-antenna UEs. To further increase the SE of the network as well as the users, the latter can be implemented with few antennas without significantly increasing the complexity and cost. Accordingly, the extension of our study to match also with multi-antenna UEs is very interesting. In this extension, it is possible to adopt at the receiver side the AD combining technique, which is the counterpart of AD-DBF and is built based on the AoA of the path having the highest power.

6.2.2 Long-term studies:

Related to the problematic considered throughout this dissertation concerning mmWave AD m-MIMO-NOMA systems, the following topics are deemed to be worthwhile pursuing as future research

- **AD m-MIMO-NOMA systems in uplink mmWave transmissions**

We have obtained significant results regarding DBF- and HBF-based schemes designed in Chapter 4 and Chapter 5, respectively, for downlink mmWave AD m-MIMO-NOMA systems. Recently, MIMO-NOMA has received attention also in uplink transmissions. However, the consideration of partial CSI was not widely considered in the literature. Therefore, there is a need to design efficient m-MIMO-NOMA systems in uplink mmWave transmissions using only angular information.

- **Effective resources allocation for HBF-based MIMO-NOMA schemes**

Chapter 5 reveals the interest of the joint SB- and MB-NOMA schemes in mmWave hybrid systems with LSAAs. However, their performance is not optimized in terms of spectral and energy efficiencies. To achieve significant gains in HBF-based MIMO-NOMA, the SB- and MB-NOMA scheme needs to be coupled with an optimal resource allocation method, including antenna selection, user clustering, and power allocation, aiming at optimizing spectral and energy efficiencies. Besides, it would be interesting to include more practical HBF structures, such as sub-connected HBF, very known for their reduced power consumption w.r.t. the fully connected HBF structure considered in Chapter 5 and allowing more energy-efficient designs.

- **AD m-MIMO-NOMA in wideband frequency-selective mmWave channels**

Since the mmWave bands are attractive due to its large available bandwidth, the wideband implementation of mmWave AD m-MIMO-NOMA systems with the orthogonal frequency division modulation (OFDM) becomes an interesting topic for next-generation cellular networks. In general, OFDM converts a frequency-selective fading channel into parallel flat fading subchannels. Therefore, the baseband precoders built on a full CSI basis involve frequency-selective designs and have unique weight vectors at each subcarrier. This leads to high hardware complexity and cost with the OFDM implementation. Interestingly, the AD-DBF beamformer only depends on the angular information of the path having the highest power, and the corresponding BF matrix is constructed based on the steering vector. Thus, AD-DBF is a frequency-flat beamformer even in a frequency-selective multipath channel. In other words, AD-DBF has the same frequency-domain response even if the propagation delay is longer than the coherence time. To this end, we expect that AD-DBF is a promising candidate over the OFDM wideband transmission reducing the hardware complexity and cost and the channel acquisition overhead. Moreover, since the mmWave channels are frequency-selective wideband channels, it would be interesting to investigate AD-DBF and the AD MIMO-NOMA schemes within OFDM transmissions and propose resource allocation with only the angular information.

Appendix A

For 2-user AD MIMO-NOMA with FCSI or ADI NOMA, the sum-throughput R_c^{type} of UEs within the c -th cluster can be expressed as

$$R_c^{\text{type}} = B \log_2(1 + \gamma_{1,c} \zeta_{1,c}^{\text{type}}) + B \log_2 \left(1 + \frac{\gamma_{2,c} \zeta_{2,c}^{\text{type}}}{\gamma_{1,c} \zeta_{2,c}^{\text{type}} + 1} \right). \quad (\text{A.1})$$

The first-order derivative of R_c^{type} w.r.t. $\gamma_{1,c} = 1 - \gamma_{2,c}$ is given by

$$\frac{dR_c^{\text{type}}}{d\gamma_{1,c}} = B \frac{\zeta_{1,c}^{\text{type}} - \zeta_{2,c}^{\text{type}}}{(1 + \zeta_{1,c}^{\text{type}} \gamma_{1,c})(1 + \zeta_{2,c}^{\text{type}} \gamma_{1,c})} \quad (\text{A.2})$$

Since $\zeta_{1,c}^{\text{type}} > \zeta_{2,c}^{\text{type}}$, then $\frac{dR_c^{\text{type}}}{d\gamma_{1,c}} > 0$ and the objective function R_c^{type} in $(\mathcal{P}_3)$ is a monotonically increasing function of $\gamma_{1,c}$. However, the constraints (4.28a) and (4.28b) in $(\mathcal{P}_3)$ implies that $0 \leq \gamma_{1,c} \leq \frac{1}{2} - \frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$. Therefore, $\gamma_{1,c}^{\star \text{type}} = \frac{1}{2} - \frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$ maximizes the sum-throughput within the 2-user NOMA cluster. This result is also obtained using KKT as in (4.37).

Bibliography

- [1] A. Osseiran, J. F. Monserrat, and P. Marsch, *5G mobile and wireless communications technology*. Cambridge University Press, 2016.
- [2] B. Tripathy and J. Anuradha, *Internet of things (IoT): technologies, applications, challenges and solutions*. CRC press, 2017.
- [3] W. Xiang, K. Zheng, and X. S. Shen, *5G mobile communications*. Springer, 2016.
- [4] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, “Five disruptive technology directions for 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.
- [5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, “What will 5G be?” *IEEE J. Sel. Areas in Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [6] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez Jr, “Millimeter wave mobile communications for 5G cellular: It will work!” *IEEE Access*, vol. 1, no. 1, pp. 335–349, 2013.
- [7] T. L. Marzetta and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge University Press, 2016.
- [8] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, “Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends,” *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, 2015.
- [9] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019, vol. 159.
- [10] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, “System-level performance evaluation of downlink non-orthogonal multiple access (NOMA),” in *Int. Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2013, pp. 611–615.
- [11] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, “Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access,” in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*. IEEE, 2013, pp. 770–774.

- [12] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2013, pp. 332–336.
- [13] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous CDMA systems over AWGN channel," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1616–1626, 2008.
- [14] Z. Wei, J. Yuan, D. W. K. Ng, M. El Kashlan, and Z. Ding, "A survey of downlink non-orthogonal multiple access for 5G wireless communication networks," *arXiv preprint arXiv:1609.01856*, 2016.
- [15] R. Razavi, A.-I. Mohammed, M. A. Imran, R. Hoshyar, and D. Chen, "On receiver design for uplink low density signature OFDM (LDS-OFDM)," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3499–3508, 2012.
- [16] S. Chen, B. Ren, Q. Gao, S. Kang, S. Sun, and K. Niu, "Pattern division multiple access—a novel nonorthogonal multiple access for fifth-generation radio networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3185–3196, 2016.
- [17] D. Zhang, Z. Zhou, C. Xu, Y. Zhang, J. Rodriguez, and T. Sato, "Capacity analysis of NOMA with mmWave massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1606–1618, 2017.
- [18] S. Wang, Y. Long, R. Ruby, and X. Fu, "Clustering and power optimization in mmwave massive MIMO–NOMA systems," *Physical Communication*, vol. 49, p. 101469, 2021.
- [19] M. Zeng, W. Hao, O. A. Dobre, and H. V. Poor, "Energy-efficient power allocation in uplink mmwave massive MIMO with NOMA," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 3000–3004, 2019.
- [20] P. Liu, Y. Li, W. Cheng, X. Gao, and W. Zhang, "Multi-beam NOMA for millimeter-wave massive MIMO with lens antenna array," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11 570–11 583, 2020.
- [21] T. S. Rappaport, S. Sun, and M. Shafi, "Investigation and comparison of 3GPP and NYUSIM channel models for 5G wireless communications," in *Proc. IEEE Vehicular Technology Conference (VTC)*, 2017.
- [22] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Tech.*, vol. 64, no. 7, pp. 2207–2225, 2016.
- [23] S. Sun, T. S. Rappaport, M. Shafi, P. Tang, J. Zhang, and P. J. Smith, "Propagation models and performance evaluation for 5G millimeter-wave bands," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8422–8439, 2018.

- [24] A. Rozé, “Massive mimo, une approche angulaire pour les futurs systèmes multi-utilisateurs aux longueurs d’onde millimétriques,” Ph.D. dissertation, INSA de Rennes, 2016.
- [25] X. Hu, C. Zhong, Y. Han, X. Chen, J. Zhao, and Z. Zhang, “Angle-domain mmWave MIMO NOMA systems: analysis and design,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019.
- [26] C. A. Balanis, *Antenna theory: analysis and design*. John Wiley & sons, 2015.
- [27] Y. Huang, *Antennas: from theory to practice*. John Wiley & Sons, 2021.
- [28] R. J. Mailloux, *Phased array antenna handbook*. Artech house, 2017.
- [29] R. C. Hansen, *Phased array antennas*. John Wiley & Sons, 2009, vol. 213.
- [30] H. J. Visser, *Array and phased array antenna basics*. Wiley Online Library, 2005.
- [31] S. Farahani, *ZigBee wireless networks and transceivers*. newnes, 2011.
- [32] 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.
- [33] W. Xiang, K. Zheng, and X. S. Shen, *5G mobile communications*. Springer, 2016.
- [34] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO wireless communications*. Cambridge university press, 2007.
- [35] N. Costa and S. Haykin, *Multiple-input multiple-output channel models: theory and practice*. John Wiley & Sons, 2010, vol. 65.
- [36] E. Telatar, “Capacity of multi-antenna gaussian channels,” *European transactions on telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.
- [37] G. G. Raleigh and J. M. Cioffi, “Spatio-temporal coding for wireless communication,” *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, 1998.
- [38] E. G. Larsson, P. Stoica, and G. Ganesan, *Space-time block coding for wireless communications*. Cambridge university press, 2003.
- [39] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [40] T. L. Marzetta and B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in rayleigh flat fading,” *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, 1999.

- [41] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, no. 6, pp. 49–83, 1997.
- [42] R. H. Roy III and B. Ottersten, "Spatial division multiple access wireless communication systems," May 7 1996, uS Patent 5,515,378.
- [43] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, 1988.
- [44] S. C. Swales, M. A. Beach, D. J. Edwards, and J. P. McGeehan, "The performance enhancement of multibeam adaptive base-station antennas for cellular land mobile radio systems," *IEEE Trans. Veh. Technol.*, vol. 39, no. 1, pp. 56–67, 1990.
- [45] T. Le-Ngoc and R. Mai, *Hybrid Massive MIMO Precoding in Cloud-RAN*. Springer, 2019.
- [46] T. L. Marzetta, "Massive MIMO: an introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015.
- [47] —, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.
- [48] —, "How much training is required for multiuser MIMO?" in *2006 fortieth asilomar conference on signals, systems and computers*. IEEE, 2006, pp. 359–363.
- [49] 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 Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2012.
- [50] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, 2013.
- [51] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive MIMO linear precoding: A survey," *IEEE Syst. J.*, no. 99, pp. 1–12, 2017.
- [52] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Processing Lett.*, vol. 21, no. 12, pp. 1501–1505, 2014.
- [53] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE Vehicular Technology Conference (VTC)*, 2013, pp. 1–5.
- [54] Q. Sun, S. Han, I. Chin-Lin, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, 2015.
- [55] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, 2015.

- [56] Z. Ding, R. Schober, and H. V. Poor, "A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4438–4454, 2016.
- [57] T. Rappaport, R. Heath, R. Daniels, and J. Murdock, *Millimeter wave wireless communications*, 2015, includes bibliographical references (pages 585–651) and index.
- [58] Y. Wang, J. Li, L. Huang, Y. Jing, A. Georgakopoulos, and P. Demestichas, "5G mobile: Spectrum broadening to higher-frequency bands to support high data rates," *IEEE Veh. Technol. Mag.*, vol. 9, no. 3, pp. 39–46, 2014.
- [59] J. I. Smith, "A computer generated multipath fading simulation for mobile radio," *IEEE Trans. Veh. Technol.*, vol. 24, no. 3, pp. 39–40, 1975.
- [60] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell system technical journal*, vol. 47, no. 6, pp. 957–1000, 1968.
- [61] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas. Propag.*, vol. 62, no. 6, pp. 3242–3256, 2014.
- [62] Y. Yu, Y. Liu, W.-J. Lu, and H.-B. Zhu, "Propagation model and channel simulator under indoor stair environment for machine-to-machine applications," in *2015 Asia-Pacific Microwave Conference (APMC)*, vol. 2. IEEE, 2015, pp. 1–3.
- [63] T. S. Rappaport, S. Y. Seidel, and K. Takamizawa, "Statistical channel impulse response models for factory and open plan building radio communicate system design," *IEEE Trans. Commun.*, vol. 39, no. 5, pp. 794–807, 1991.
- [64] Wireless Valley Communications, Inc., *SMRCIM Plus 4.0 (Simulation of Mobile Radio Channel Impulse Response Models) User's Manual*, 1999.
- [65] P. Kyösti, J. Meinilä, L. Hentilä, X. Zhao, T. Jämsä, C. Schneider, M. Narandzić, M. Milojević, A. Hong, J. Ylitalo *et al.*, "WINNER II D1. 1.2 V1. 2 WINNER II channel models," *Inf. Soc. Technol., Nokia Siemens Netw. GMBH & CO. KG, Munich, Germany, Tech. Rep. IST-4-027756*, 2007.
- [66] B. Mondal, T. A. Thomas, E. Visotsky, F. W. Vook, A. Ghosh, Y.-H. Nam, Y. Li, J. Zhang, M. Zhang, Q. Luo *et al.*, "3D channel model in 3GPP," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 16–23, 2015.
- [67] Y. He, Y. Zhang, J. Zhang, L. Pang, Y. Chen, and G. Ren, "Investigation and comparison of QuaDRiGa, NYUSIM and MG5G channel models for 5G wireless communications," in *2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall)*. IEEE, 2020.

- [68] T. S. Rappaport, G. R. MacCartney, S. Sun, H. Yan, and S. Deng, “Small-scale, local area, and transitional millimeter wave propagation for 5G communications,” *IEEE Trans. Antennas. Propag.*, vol. 65, no. 12, pp. 6474–6490, 2017.
- [69] 3GPP, “Study on channel model for frequencies from 0.5 to 100 GHz,” *TR 38.901 V14.3.0, 3rd Generation Partnership Project (3GPP)*, Dec 2017.
- [70] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, “Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design,” *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3029–3056, 2015.
- [71] M. K. Samimi and T. S. Rappaport, “Local multipath model parameters for generating 5G millimeter-wave 3GPP-like channel impulse response,” in *2016 10th European Conference on Antennas and Propagation (EuCAP)*. IEEE, 2016.
- [72] G. S. Smith, “A direct derivation of a single-antenna reciprocity relation for the time domain,” *IEEE Trans. Antennas. Propag.*
- [73] J. Choi, D. J. Love, and P. Bidigare, “Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, 2014.
- [74] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, “Pilot contamination and precoding in multi-cell TDD systems,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, 2011.
- [75] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, “A coordinated approach to channel estimation in large-scale multiple-antenna systems,” *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 2, pp. 264–273, 2013.
- [76] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, “Blind pilot decontamination,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014.
- [77] Y. Wu, T. Liu, M. Cao, L. Li, and W. Xu, “Pilot contamination reduction in massive MIMO systems based on pilot scheduling,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 1–9, 2018.
- [78] J.-C. Guey and L. D. Larsson, “Modeling and evaluation of MIMO systems exploiting channel reciprocity in TDD mode,” in *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, vol. 6. IEEE, 2004, pp. 4265–4269.
- [79] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, “Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

- [80] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 711–723, 2014.
- [81] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2868–2882, 2015.
- [82] X. Rao and V. K. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Processing*, vol. 62, no. 12, pp. 3261–3271, 2014.
- [83] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Processing*, vol. 63, no. 23, pp. 6169–6183, 2015.
- [84] B. Zhang, L. J. Cimini, and L. J. Greenstein, "Efficient eigenspace training and precoding for FDD massive MIMO systems," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.
- [85] B. Zhang, G. Yue, and L. J. Cimini, "Efficient beamforming training and limited feedback precoding for massive MIMO systems," *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 9, pp. 2197–2214, 2020.
- [86] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE Access*, vol. 5, pp. 7667–7681, 2017.
- [87] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.
- [88] M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [89] C. Windpassinger, R. F. Fischer, and J. B. Huber, "Lattice-reduction-aided broadcast precoding," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2057–2060, 2004.
- [90] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part ii: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537–544, 2005.
- [91] S. Dutta, C. N. Barati, A. Dhananjay, D. A. Ramirez, J. F. Buckwalter, and S. Rangan, "A case for digital beamforming at mmWave," *arXiv preprint arXiv:1901.08693*, 2019.
- [92] A. Rozé, M. H  lard, M. Cruss  re, and C. Langlais, "Millimeter-wave digital beamsteering in highly line-of-sight environments for massive MIMO systems," in *Wireless World Research Forum Meeting*, vol. 35, 2015.

- [93] —, “Linear precoder performance for massive MIMO systems in near LoS environments: Application to mmwave transmission,” in *Proceedings of European Wireless 2015; 21th European Wireless Conference*. VDE, 2015.
- [94] I. Khaled, A. El Falou, C. Langlais, B. ElHassan, and M. Jezequel, “Multi-user digital beamforming based on path angle information for mm-Wave MIMO systems,” in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*. VDE, 2020.
- [95] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, “Channel estimation and hybrid precoding for millimeter wave cellular systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.
- [96] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, “Millimeter wave beamforming for wireless backhaul and access in small cell networks,” *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, 2013.
- [97] I. Khaled, A. EL Falou, C. Langlais, B. ELHassan, and M. Jezequel, “Performance evaluation of linear precoding mmWave multi-user MIMO systems with NYUSIM channel simulator,” in *Proc. IEEE Middle East and North Africa COMMUNICATIONS Conference (MENACOMM)*, 2019.
- [98] Z. Ma, Z. Zhang, Z. Ding, P. Fan, and H. Li, “Key techniques for 5g wireless communications: network architecture, physical layer, and mac layer perspectives,” *Science China information sciences*, vol. 58, no. 4, pp. 1–20, 2015.
- [99] I. Khaled, C. Langlais, A. El Falou, M. Jezequel, and B. ElHassan, “Joint SDMA and power-domain NOMA system for multi-user mm-Wave communications,” in *Proc. IEEE Int. Wireless Commun. and Mobile Computing (IWCMC)*, 2020, pp. 1112–1117.
- [100] I. Khaled, C. Langlais, A. El Falou, B. ElHassan, and M. Jezequel, “Multi-user angle-domain MIMO-NOMA system for mmWave communications,” *IEEE Access*, 2021.
- [101] S. Liu, C. Zhang, and G. Lyu, “User selection and power schedule for downlink non-orthogonal multiple access (NOMA) system,” in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, 2015, pp. 2561–2565.
- [102] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.-G. Xia, “Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961–2974, 2018.
- [103] X. Chen, F.-K. Gong, G. Li, H. Zhang, and P. Song, “User pairing and pair scheduling in massive MIMO-NOMA systems,” *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 788–791, 2017.
- [104] T.-V. Nguyen, V.-D. Nguyen, D. B. da Costa, and B. An, “Hybrid user pairing for spectral and energy efficiencies in multiuser MISO-NOMA networks with SWIPT,” *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4874–4890, 2020.

- [105] A. S. de Sena, F. R. M. Lima, D. B. da Costa, Z. Ding, P. H. Nardelli, U. S. Dias, and C. B. Papadias, “Massive MIMO-NOMA networks with imperfect SIC: Design and fairness enhancement,” *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6100–6115, 2020.
- [106] X. Zhang, X. Zhu, and H. Zhu, “Joint user clustering and multi-dimensional resource allocation in downlink MIMO-NOMA networks,” *IEEE Access*, vol. 7, pp. 81 783–81 793, 2019.
- [107] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, “Energy-efficient power allocation for MIMO-NOMA with multiple users in a cluster,” *IEEE Access*, vol. 6, pp. 5170–5181, 2018.
- [108] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, “Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array,” *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 10, pp. 2370–2382, 2017.
- [109] J. Choi, “Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems,” *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, 2015.
- [110] J. Wang, S. Jin, X. Gao, K.-K. Wong, and E. Au, “Statistical eigenmode-based SDMA for two-user downlink,” *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5371–5383, 2012.
- [111] Y. Sun, J. Zhou, Q. Cao, and S. Li, “Precoder design in statistical CSI aided non-orthogonal multiple access,” *IEEE Access*, vol. 6, pp. 16 484–16 492, 2018.
- [112] J. Choi, “Joint rate and power allocation for NOMA with statistical CSI,” *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4519–4528, 2017.
- [113] M. M. Al-Wani, A. Sali, N. K. Noordin, S. J. Hashim, C. Y. Leow, and I. Krikidis, “Robust beamforming and user clustering for guaranteed fairness in downlink NOMA with partial feedback,” *IEEE Access*, vol. 7, pp. 121 599–121 611, 2019.
- [114] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, “Overview of millimeter wave communications for fifth-generation (5G) wireless networks—with a focus on propagation models,” *IEEE Trans. Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.
- [115] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, “Millimeter wave channel modeling and cellular capacity evaluation,” *IEEE J. Sel. Areas in Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.
- [116] J. Wang, Y. Li, C. Ji, Q. Sun, S. Jin, and T. Q. Quek, “Location-based MIMO-NOMA: Multiple access regions and low-complexity user pairing,” *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2293–2307, 2020.

- [117] E. M. Mohamed, “Joint users selection and beamforming in downlink millimetre-wave NOMA based on users positioning,” *IET Communications*, vol. 14, no. 8, pp. 1234–1240, 2020.
- [118] X. Lu, Y. Zhou, and V. W. Wong, “A joint angle and distance based user pairing strategy for millimeter wave NOMA networks,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2020.
- [119] Z. Zheng, W.-Q. Wang, H. Meng, H. C. So, and H. Zhang, “Efficient beamspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 776–11 789, 2018.
- [120] T. Wang, B. Ai, R. He, and Z. Zhong, “Two-dimension direction-of-arrival estimation for massive MIMO systems,” *IEEE Access*, vol. 3, pp. 2122–2128, 2015.
- [121] M. S. Ali, H. Tabassum, and E. Hossain, “Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems,” *IEEE Access*, vol. 4, pp. 6325–6343, 2016.
- [122] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, “38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2012, pp. 4568–4573.
- [123] N. Rupasinghe, Y. Yapıcı, I. Güvenç, M. Ghosh, and Y. Kakishima, “Angle feedback for NOMA transmission in mmWave drone networks,” *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 628–643, 2019.
- [124] Q. Yang, H.-M. Wang, D. W. K. Ng, and M. H. Lee, “NOMA in downlink SDMA with limited feedback: Performance analysis and optimization,” *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 10, pp. 2281–2294, 2017.
- [125] M. Morales-Céspedes, O. A. Dobre, and A. García-Armada, “Semi-blind interference aligned NOMA for downlink MU-MISO systems,” *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1852–1865, 2019.
- [126] Q. N. Le, V.-D. Nguyen, N.-P. Nguyen, S. Chatzinotas, O. A. Dobre, and R. Zhao, “Learning-assisted user clustering in cell-free massive MIMO-NOMA networks,” *arXiv preprint arXiv:2011.07549*, 2020.
- [127] S. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, “Resource allocation for downlink NOMA systems: Key techniques and open issues,” *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 40–47, 2018.
- [128] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir, “Unsupervised machine learning-based user clustering in millimeter-wave-NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7425–7440, 2018.

- [129] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, "Millimeter-wave NOMA with user grouping, power allocation and hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5065–5079, 2019.
- [130] E. K. Chong and S. H. Zak, *An introduction to optimization*. John Wiley & Sons, 2004.
- [131] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal multiple access with SIC in cellular downlink using proportional fair-based resource allocation," *IEICE Trans. Commun.*, vol. 98, no. 2, pp. 344–351, 2015.
- [132] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, 2017.
- [133] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, 2018.
- [134] J. Zhang, L. Dai, S. Sun, and Z. Wang, "On the spectral efficiency of massive MIMO systems with low-resolution ADCs," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 842–845, 2016.
- [135] J. Zhang, L. Dai, Z. He, S. Jin, and X. Li, "Performance analysis of mixed-ADC massive MIMO systems over rician fading channels," *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 6, pp. 1327–1338, 2017.
- [136] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, 2014.
- [137] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas in Commun.*, vol. 34, no. 4, pp. 998–1009, 2016.
- [138] C. Lin and G. Y. Li, "Energy-efficient design of indoor mmWave and sub-THz systems with antenna arrays," *IEEE Trans. Commun.*, vol. 15, no. 7, pp. 4660–4672, 2016.
- [139] I. Khaled, A. El Falou, C. Langlais, B. El Hassan, and M. Jezequel, "Joint single-and multi-beam angle-domain NOMA for hybrid mmWave MIMO systems," in *WSA 2021; 25th International ITG Workshop on Smart Antennas*. VDE, 2021.
- [140] W. Yuan, V. Kalokidou, S. M. Armour, A. Doufexi, and M. A. Beach, "Application of non-orthogonal multiplexing to mmWave multi-user systems," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, 2017, pp. 1–6.

- [141] X. Yu, F. Xu, K. Yu, and X. Dang, "Power allocation for energy efficiency optimization in multi-user mmWave-NOMA system with hybrid precoding," *IEEE Access*, vol. 7, pp. 109 083–109 093, 2019.
- [142] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas in Commun.*, vol. 37, no. 1, pp. 131–141, 2018.
- [143] A. N. Uwaechia and N. M. Mahyuddin, "Spectrum and energy efficiency optimization for hybrid precoding-based SWIPT-enabled mmWave mMIMO-NOMA systems," *IEEE Access*, vol. 8, pp. 139 994–140 007, 2020.
- [144] A. A. Badrudeen, C. Y. Leow, and S. Won, "Performance analysis of hybrid beamforming precoders for multiuser millimeter wave NOMA systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8739–8752, 2020.
- [145] L. Pang, W. Wu, Y. Zhang, Y. Yuan, Y. Chen, A. Wang, and J. Li, "Joint power allocation and hybrid beamforming for downlink mmwave-NOMA systems," *IEEE Trans. Veh. Technol.*, 2021.
- [146] A. A. Badrudeen, C. Y. Leow, and S. Won, "Sub-connected structure hybrid precoding for millimeter-wave NOMA communications," *IEEE Wireless Commun. Lett.*, vol. 10, no. 6, pp. 1334–1338, 2021.
- [147] Z. Xiao, L. Dai, Z. Ding, J. Choi, and P. Xia, "Millimeter-wave communication with non-orthogonal multiple access for 5G," *arXiv preprint arXiv:1709.07980*, 2017.
- [148] Z. Wei, L. Zhao, J. Guo, D. W. K. Ng, and J. Yuan, "A multi-beam NOMA framework for hybrid mmWave systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*. IEEE, 2018.
- [149] —, "Multi-beam NOMA for hybrid mmWave systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1705–1719, 2018.
- [150] X. Su, J. Zeng, J. Li, L. Rong, L. Liu, X. Xu, and J. Wang, "Limited feedback precoding for massive MIMO," *International Journal of Antennas and Propagation*, vol. 2013, 2013.
- [151] G. Lee, Y. Sung, and J. Seo, "Randomly-directional beamforming in millimeter-wave multiuser MISO downlink," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1086–1100, 2015.

Titre : Formation de voie multi-utilisateurs dans le domaine angulaire et accès multiple non orthogonal pour les systèmes multi-antennes massifs en ondes millimétriques

Mots clés : transmissions sans fil, ondes millimétriques, antennes multiples, formation de voies, précodage, accès multiple, techniques multi-utilisateurs, connaissance partielle du canal

Résumé : Pour répondre à la croissance rapide du nombre d'appareils connectés et aux exigences de qualité de service des services innovants, les générations cellulaires 5G et suivantes proposent de nouvelles technologies, telles que les bandes à ondes millimétriques (mmWave), les systèmes multi-antennes massives (MIMO) et l'accès multiple non orthogonal (NOMA). L'intégration de ces technologies peut améliorer considérablement la capacité du réseau. Cependant, la mise en œuvre d'un nombre massif d'antennes dans les bandes d'ondes millimétriques pose de nombreux problèmes, notamment en termes de consommation d'énergie, de complexité matérielle et de surcharge d'estimation des canaux. L'objectif de ce travail est de résoudre ces problèmes dans les systèmes MIMO-NOMA massifs à ondes millimétriques. Pour ce faire, nous exploitons la caractéristique clé des canaux à ondes millimétriques, c'est-à-dire la transmission fortement directionnelle, qui fait de l'information angulaire une information partielle prometteuse sur l'état du canal. Tout d'abord, nous étudions les performances d'un précodeur de formation de faisceau numérique dans le domaine angulaire (AD-DBF) peu complexe en utilisant le modèle de canal à ondes millimétriques réaliste et statistique appelé NYUSIM de l'université de New York. Ensuite, nous proposons un schéma MIMO-NOMA dans le domaine angulaire pour permettre de desservir un plus grand nombre d'utilisateurs que le schéma AD-DBF classique. En particulier, nous dérivons de nouvelles métriques basées sur l'angle pour concevoir des techniques de regroupement et d'ordonnement des utilisateurs dans le domaine angulaire et d'allocation de puissance. Troisièmement, nous adoptons la formation de faisceau hybride dans le domaine angulaire pour réduire la consommation d'énergie, la complexité et le coût, lorsque des centaines d'antennes (par exemple, 128, 256) sont mises en œuvre au niveau de l'émetteur. Plus précisément, nous proposons différents schémas de domaine angulaire qui visent à exploiter davantage de degrés de liberté et à permettre la communication même lorsque le nombre d'utilisateurs est supérieur à celui des chaînes de radiofréquence. Dans ce travail, nous avons montré que cette communication MIMO-NOMA massive en ondes millimétriques basée sur l'information angulaire permet une connectivité massive des utilisateurs, même dans des scénarios surchargés, avec une faible surcharge d'estimation de canal. De plus, pour réduire la complexité et le coût, une version hybride est disponible.

Title : Multi-user angle-domain beamforming and non-orthogonal multiple access for millimeter-wave massive MIMO systems

Keywords : wireless transmissions, millimeter waves, multi-antennas, beamforming, precoding, multiple access, multi-user techniques, partial channel state information

Abstract : To meet the rapid growth in the number of connected devices and the quality-of-service requirements of innovative services, the 5G and beyond cellular generations have suggested new technologies, such as millimeter-wave (mmWave) bands, massive multiple-input multiple-output (MIMO), and non-orthogonal multiple access (NOMA). The integration of these technologies, namely "massive MIMO-NOMA-aided mmWave communications," can significantly enhance the network capacity. However, the implementation of a massive number of antennas at mmWave bands poses many challenging issues, including power consumption, hardware complexity, and channel estimation overhead. The aim of this work is to tackle these issues in the mmWave massive MIMO-NOMA systems. To do so, we exploit the key feature of mmWave channels, i.e., the highly directional transmission, that makes the angular information a promising partial channel state information. First, we investigate the performance of a low-complex angle-domain digital beamforming (AD-DBF) precoder in mmWave massive-MIMO systems using the realistic and statistical mmWave channel model called NYUSIM from New York University. Second, we propose an angle-domain MIMO-NOMA scheme to enable the capacity of serving more users compared to the conventional AD-DBF. In particular, we derive new angular-based metrics to design angle-domain user clustering and ordering, and power allocation techniques. Third, we adopt angle-domain hybrid beamforming to reduce power consumption, complexity and cost, when hundreds of antennas (e.g., 128, 256) are implemented at the transmitter. Specifically, we propose different angle-domain schemes that aim at leveraging more degrees of freedom and enabling communication even when the number of users is greater than that of radio-frequency chains. In this work, we have shown that this massive MIMO-NOMA mmWave communication based on angular information provides massive user connectivity, even in overloaded scenarios, with low channel estimation overhead. Moreover, to lower complexity and cost, a hybrid version is available.