UNIVERSITÀ DEGLI STUDI DI CASSINO E DEL LAZIO MERIDIONALE

DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELL'INFORMAZIONE

Massive MIMO Technologies for 5G and Beyond-5G Wireless Networks

---

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Electrical and Information Engineering

January 2019

RESEARCH SUPERVISORS:
Prof. Stefano BUZZI
Dr. Alessio ZAPPONE

AUTHOR:
Carmen D'ANDREA

Abstract

Massive multiple input multiple output (MIMO) is a promising 5G and beyond-5G wireless access technology that can provide huge throughput, compared with the current technology, in order to satisfy some requirements for the future generations of wireless networks. The research described in this thesis proposes the design of some applications of the massive MIMO technology that can be implemented in order to increase the spectral efficiency per cell of the future wireless networks through a simple and low complexity signal processing. In particular, massive MIMO is studied in conjunction with two other topics that are currently under investigation for the future wireless systems, both in academia and in industry: the millimeter wave frequencies and the distributed antenna systems. The first part of the thesis gives a brief overview on the requirements of the future wireless networks and it explains some of the mathematical tools used in the current massive MIMO literature. Then, an overview on the differences between massive MIMO techniques at the conventional cellular frequencies and at millimeter wave frequencies is presented and exhaustively discussed. Six key basic differences are pinpointed, along with the implications that they have on the architecture and algorithms of the communication transceivers and on the attainable performance in terms of reliability and multiplexing capabilities. Subsequently, "doubly massive MIMO" systems at millimeter wave frequencies are introduced, i.e., systems with a large number of antennas at both the transmitter and the receiver. For complexity reasons and energy consumption issues, fully digital pre-coding and post-coding structures may turn out to be unfeasible, and thus suboptimal structures, making use of simplified hardware and a limited number of radio-frequency chains, have been investigated. A comparative assessment of several suboptimal pre-coding and post-coding structures with large number of antennas is discussed. Numerical results show that fully-digital beamformers may actually achieve a larger energy efficiency than lower-complexity solutions, as well as that low-complexity beam-steering purely analog beamforming may in some cases represent a good performance-complexity trade-off solution. Finally, the thesis focuses on the recently introduced cell-free massive MIMO architecture, wherein a very large number of distributed access points, connected to a central processing unit, simultaneously and jointly serve a much smaller number of mobile stations. It contrasts the originally proposed formulation of cell-free massive MIMO with a user-centric approach wherein each mobile station is served only by a limited number of access points. Exploiting the framework of successive lower-bound maximization, this thesis also proposes and analyzes two power allocation strategies aimed at maximizing the throughput and the fairness of these systems. Additionally, advanced signal processing techniques, to improve the performance of the user-centric approach both in uplink and in downlink, are proposed. The proposed schemes can be implemented locally, i.e., with no need to exchange information with the central

processing unit. Numerical results show that the user-centric approach, which requires smaller backhaul overhead and it is more scalable than the cell-free massive MIMO deployment, also achieves generally better performance than the cell-free massive MIMO approach for the vast majority of the users in the system, especially on the uplink. Regarding the proposed advanced signal processing techniques, the results show that they provide remarkable performance improvements with respect to the competing alternatives.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC	**A**nalog-to-**D**igital **C**onverter
AN	**AN**alog
AP	**A**ccess **P**oint
ASE	**A**chievable **S**pectral **E**fficiency
BS	**B**ase **S**tation
CF	**C**ell-**F**ree
CM	**C**hannel **M**atched
CM-FD	**C**hannel **M**atched, **F**ully **D**igital
CM-HY	**C**hannel **M**atched, **HY**brid
CoMP	**Co**ordinated **M**ulti-**P**oint
CPU	**C**entral **P**rocessing **U**nit
CSI	**C**hannel **S**tate **I**nformation
DAC	**D**igital-to-**A**nalog-**C**onverter
EM	**E**lectro**M**agnetic
eMBB	**E**nhanced **M**obile **B**roadband
FBWA	**F**ixed **B**roadband **W**ireless **A**ccess
FD	**F**ully **D**igital
FDD	**F**requency **D**ivision **D**uplex
FFT	**F**ast **F**ourier **T**ransform
GEE	**G**lobal **E**nergy **E**fficiency
HY	**HY**brid
i.i.d.	**i**ndipendent and **i**dentical **d**istributed
InH Office	**In**door **H**otspot **Office**
InH Shopping Mall	**In**door **H**otspot **Shopping Mall**
KKT	**K**arush–**K**uhn–**T**ucker
LOS	**L**ine **O**f **S**ight
LNA	**L**ow **N**oise **A**mplifier
LTE	**L**ong **T**erm **E**volution
MFN	**M**inimum **F**robenius **N**orm
MIMO	**M**ultiple **I**nput **M**ultiple **O**utput
MMSE	**M**inimum **M**ean **S**quare **E**rror
mm-Wave	Millimeter (**mm**)-**Wave**
MR	**M**aximum **R**atio
MTC	**M**achine **T**ype **C**ommunication
MU-MIMO	**MU**ltiuser **MIMO**

MS	**M**obile **S**tation
NLMS	**N**ormalized **L**east-**M**ean-**S**quares
NLOS	**N**on **L**ine **O**f **S**ight
NOMA	**N**on-**O**rthogonal **M**ultiple **A**ccess
OFDM	**O**rthogonal **F**requency **D**ivision **M**ultiplexing
OOJA	**O**rthogonal **OJA**
PA	**P**ower **A**mplifier
PAPR	**P**eak-to-**A**verage **P**ower **R**atio
PASTd	**P**rojection **A**pproximation **S**ubspace **T**racking with **d**eflation
PM	**P**ilot **M**atched
PZF	**P**artial **Z**ero **F**orcing
PZF-FD	**P**artial **Z**ero **F**orcing, **F**ully **D**igital
PZF-HY	**P**artial **Z**ero **F**orcing, **HY**brid
RF	**R**adio **F**requency
RLS	**R**ecursive-**L**east-**S**quares
RV	**R**andom **V**ariable
SCM	**S**ingle-**C**arrier **M**odulation
SDMA	**S**pace **D**ivision **M**ultiple **A**ccess
SIC	**S**uccessive **I**nterference **C**ancellation
SINR	**S**ignal-to-**I**nterference plus **N**oise **R**atio
SNR	**S**ignal to **N**oise **R**atio
SU-MIMO	**S**ingle **U**ser **MIMO**
SVD	**S**ingular-**V**alue-**D**ecomposition
SW	**SW**itches
SW+PHSH	**SW**itches and fixed **PH**ase **SH**ifters
TDD	**T**ime **D**ivision **D**uplex
TIN	**T**reating **I**nterference as **N**oise
UatF	**U**se-**a**nd-**t**hen-**F**orget
UAV	**U**nmanned **A**erial **V**ehicles
UC	**U**ser **C**entric
UFMC	**U**niversal **F**iltered **M**ulti**C**arrier
ULA	**U**niform **L**inear **A**rray
UMi Open-Square	**U**rban **Mi**crocellular **Open-Square**
UMi Street-Canyon	**U**rban **Mi**crocellular **Street-Canyon**
URLLC	**U**ltra-**R**eliable **L**ow-**L**atency **C**ommunications
V2I	**V**ehicle-to-**I**nfrastructure
V2V	**V**ehicle-to-**V**ehicle
ZF	**Z**ero **F**orcing
5G	Fifth (**5**th) **G**eneration
μ-Wave	Micro (μ)-**Wave**
3GPP	Third (**3**rd) **G**eneration **P**artnership **P**roject
5G-NR	**5G**-**N**ew **R**adio

Mathematical Notation

$\mathbf{A}$	generic matrix
$\mathbf{a}$	generic vector
a	generic scalar
$\mathbb{C}^{N\times M}$	the set of complex-valued $N \times M$ matrices, when $M = 1$ or $N = 1$ they are omitted
$\mathbb{R}^{N\times M}$	the set of real-valued $N \times M$ matrices, when $M = 1$ or $N = 1$ they are omitted
$a \in \mathcal{S}$	a is a member of the set $\mathcal{S}$
$a \notin \mathcal{S}$	a is not a member of the set $\mathcal{S}$
card $[\mathcal{S}]$	the cardinality of the set $\mathcal{S}$
$\mathbf{A}^*$	conjugate of matrix $\mathbf{A}$
$\mathbf{A}^T$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of matrix $\mathbf{A}$
$\mathbf{A}^H$	conjugate transpose of matrix $\mathbf{A}$
$\mathbf{A}^+$	pseudo-inverse of matrix $\mathbf{A}$
$\lvert\mathbf{A}\rvert$	determinant of matrix $\mathbf{A}$
$\lvert a\rvert$	absolute value of the scalar a
$\underline{/a}$	the phase of the complex scalar a
j	imaginary unit
e	Euler's number ($e \approx 2.718281828$)
$(\mathbf{a})_i$	the $i-$th element of a vector $\mathbf{a}$
$(\mathbf{A})_{(i,\ell)}$	the $(i,\ell)-$th element of a matrix $\mathbf{A}$
$(\mathbf{A})_{(i,:)}$	the $i-$th row of a matrix $\mathbf{A}$
$(\mathbf{A})_{(:,\ell)}$	the $\ell-$th column of a matrix $\mathbf{A}$
$(\mathbf{A})_{(i:i',\ell:\ell')}$	matrix that contains rows from the $i-$th to the i'-th and columns from the $\ell-$th to the ℓ'-th of matrix $\mathbf{A}$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of $\mathbf{A}, \mathbf{B}$
$\mathrm{tr}(\mathbf{A})$	trace of a matrix $\mathbf{A}$
$\mathcal{N}(\mathbf{a}, \mathbf{R})$	real Gaussian distribution with mean $\mathbf{a}$ and correlation matrix $\mathbf{R}$
$\mathcal{CN}(\mathbf{a}, \mathbf{R})$	complex Gaussian distribution with mean $\mathbf{a}$ and correlation matrix $\mathbf{R}$
$\log_b(\cdot)$	logarithm function using the base b

$\sin(\cdot)$	sine function
$\cos(\cdot)$	cosine function
$\min(a, b)$	the minimum of a and b
$\max(a, b)$	the maximum of a and b
$\|\mathbf{a}\|$	the L2-norm of a vector $\mathbf{a}$ $\|\mathbf{a}\| = \sqrt{\sum_i [(\mathbf{a})_i]^2}$
$\|\mathbf{A}\|_F$	The Frobenius norm of a matrix $\mathbf{A}$ $\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} \left[(\mathbf{A})_{(i,j)}\right]^2}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbb{E}[\cdot]$	statistical expectation operator
$\mathrm{Var}[\cdot]$	variance operator
$\mathbf{1}_{N\times M}$	the $N \times M$ matrix with only ones, when $M = 1$ or $N = 1$ they are omitted
$\mathbf{0}_{N\times M}$	the $N \times M$ matrix with only zeros, when $M = 1$ or $N = 1$ they are omitted

A tutte le persone che hanno creduto in me.

To all the people who believed in me.

Chapter 1

Introduction

1.1 Future wireless networks and massive MIMO

The term "5G and beyond-5G wireless networks" refers to future generations of mobile wireless communication systems. The vision for these next-generation systems is to enable groundbreaking mobile applications requiring high-quality low-latency visual, tactile and audio telepresence, in addition to massive capacity (upwards of 1000 times with respect to 4G systems) and connectivity (billions of users and machines) [1]. Of these requirements, certainly the one that gets the most attention is the need for radically higher data rates across the board, because current coding techniques are very close to the theoretical Shannon spectral efficiency bounds limits for single user capacity. For this reason, it is necessary to introduce new technologies in the design of future wireless networks. Wireless communication is based on radio, meaning that electromagnetic (EM) waves are designed to carry information from a transmitter to one or more receivers. Since the EM waves propagate in all possible directions from the transmitter, the signal energy spreads out and less energy reaches a desired receiver as the distance increases. To deliver wireless services with sufficiently high received signal energy over wide coverage areas, researchers at Bell Labs proposed in 1947 to deploy a cellular network topology [2]. According to this idea, the coverage area is divided into cells that operate individually using a fixed-location base station (BS); that is, a piece of network equipment that facilitates wireless communication between a device and the network. The cellular concept was further developed and analyzed over the subsequent decades and later deployed in practice. Without any doubt, the cellular concept was a major breakthrough and has been the main driver to deliver wireless services in the last forty years (since the "first generation" of mobile phone systems emerged in the 1980s). Briefly, a cellular network consists of a set of BSs and a set of mobile stations (MSs), also simply referred to as users. Each MS is connected to one of the BSs, which provides service to it. The downlink refers to signals sent from the BSs to their respective MSs, while the uplink refers to transmissions from the MSs to their respective BSs. Cellular networks were originally designed for wireless voice communications, but it is wireless data transmission

that dominates nowadays. Video on-demand accounts for the majority of traffic in wireless networks and is also the main driver of the predicted increase in traffic demand. The *area throughput* is thus a highly relevant performance metric of contemporary and future cellular networks. It is measured in $\mathrm{bit/s/m^2}$ and can be modeled using the following high-level formula:

$$\text{Area throughput} \left[\frac{\text{bit}}{\text{s m}^2}\right] = B\,[\text{Hz}]\; D \left[\frac{\text{cell}}{\text{m}^2}\right] \text{SE} \left[\frac{\text{bit}}{\text{s Hz cell}}\right], \tag{1.1}$$

where B is the bandwidth, D is the average cell density, and SE is the spectral efficiency per cell, i.e., the amount of information that can be transferred per second over one Hz of bandwidth in each cell. These are the three main components that determine the area throughput, and that need to be increased in order to achieve higher area throughput in future cellular networks. It is possible to obtain the $1000\times$ requirement in the capacity of the network as [1]

$$\underbrace{\text{Area throughput}}_{1000\times} = \underbrace{B}_{10\times}\ \underbrace{D}_{10\times}\ \underbrace{\text{SE}}_{10\times}. \tag{1.2}$$

The massive capacity in the future wireless networks can be obtained designing a network using the three new technologies detailed in the following:

(a) Extreme densification and offloading to improve the density of cells per area, i.e., more cells per unit area.

(b) Increased bandwidth, primarily by moving toward and into millimeter wave (mm-Wave) spectrum but also by making better use of WiFi's unlicensed spectrum in the 5-GHz band, i.e., more Hz.

(c) Increased spectral efficiency per cell, primarily through advances in MIMO, such as massive MIMO.

This thesis focuses on the design of some **massive MIMO technologies** that can be implemented in order to increase the spectral efficiency per cell of the future wireless networks. In particular, in the following chapters, massive MIMO is used in conjunction with two other current research topics for the future wireless networks: the mm-Wave frequencies and the distributed antenna systems.
Stemming from research that blossomed in the late 1990s, MIMO communication was introduced into WiFi systems around 2006 and into 3G cellular shortly thereafter. In essence, MIMO embodies the spatial dimension of the communication that arises once a multiplicity of antennas are available at BSs and mobile devices. If the entries of the channel matrix exhibit sufficient statistical independence by virtue of spacing, cross-polarization and/or angular disposition multiple spatial dimensions become available for signaling and the spectral efficiency increases, as explained for example in [3]. In single-user MIMO (SU-MIMO), the dimensions are limited by the number of antennas that can be accommodated on a mobile device, because of limited space. However, by having

each BS communicating with several users concurrently, the multiuser version of MIMO (MU-MIMO) can effectively aggregate the users antennas and overcome this bottleneck. Then, the signaling dimensions are given by the smallest between the aggregate number of antennas at those users and the number of antennas at the BS. Furthermore, in what is now known as coordinated multipoint (CoMP) transmission/reception, multiple BSs can cooperate and act as a single effective MIMO transceiver thereby turning some of the interference in the system into useful signals; this concept in fact underpins many of the approaches to interference and mobility management. Well-established by the time long term evolution (LTE) was developed, MIMO was a native ingredient thereof with two-to-four antennas per mobile device and as many as eight per BS sector, and it appeared that, because of form factors and other apparent limitations, such was the extent to which MIMO could be leveraged.
Marzetta was instrumental in articulating a vision in which the number of antennas increased by more than an order of magnitude, with the landmark paper [4]. The proposal was to equip BSs with a number of antennas much larger than the number of active users per time–frequency signaling resource, and given that under reasonable time–frequency selectivities accurate channel estimation can be conducted for at most some tens of users per resource, this condition puts hundreds of antennas per BS. This bold idea, initially termed "large-scale antenna systems" but now more popularly known as "massive MIMO", offers enticing benefits:

- *Enormous enhancements in spectral efficiency* without the need for increased BS densification, with the possibility of trading some of those enhancements off for power efficiency improvements [5], [6].

- *Smoothed out channel responses* because of the vast spatial diversity, which brings about the favorable action of the law of large numbers. In essence, all small-scale randomness abates as the number of channel observations grows.

- *Simple transmit/receive structures* because of the quasi-orthogonal nature of the channels between each BS and the set of active users sharing the same signaling resource. For a given number of active users, such orthogonality sharpens as the number of BS antennas grows and even plain single-user beamforming, perform close-to-optimally. In a practical multicell system, these benefits are not actually obtained, due to the effect of pilot contamination, and the finit number of BS antennas.

The promise of these benefits has elevated the massive MIMO technology to a central position in the research and developments on 5G and beyond-5G wireless networks [7], with a foreseen role of providing a high-capacity umbrella of

ubiquitous coverage in support of underlying tiers of small cells. However, despite these significant benefits, several challenges have begun to be addressed in the recent past both in academia and industry. In particular:

1. *Pilot contamination and overhead reduction*: Pilot transmissions can be made orthogonal among same-cell users, to facilitate cleaner channel estimates [8], [9] but must be reused across cells, otherwise all available resources would be consumed by pilots. This inevitably causes interference among pilots in different cells and hence puts a floor on the quality of the channel estimates. This, so-called "pilot contamination" interference, does not vanish as the number of BS antennas grows large. However, it was recenly shown that, using more realistic channel models and high performing signal processing schemes, the spectral efficiency of the system grows unbounded as the number of antennas increases, even under pilot contamination [10], [11].

2. *Architectural challenges*: A more serious challenge to the realization of the massive MIMO has to do with its architecture. This vision requires radically different BS structures where, in lieu of a few high-power amplifiers feeding a handful of sector antennas, a myriad of tiny antennas fed by correspondingly low-power amplifiers are needed; most likely each antenna would have to be integrated with its own amplifier. Scalability, antenna correlations, mutual couplings and cost are some of the issues that must be sorted out. At the same time, opportunities arise for innovative topologies such as conformal arrays along rooftops or on building facades. As an example, references [12], [13] explore alternative and highly innovative antenna designs based on the utilization of an electromagnetic lens-focusing antenna.

3. *Full-Dimension MIMO and Elevation Beamforming*: Existing BSs mostly feature linear horizontal arrays, which in tower structures can only accommodate a limited number of antennas, due to form factors, and which only exploit the azimuth angle dimension. By adopting planar 2D arrays and further exploiting the elevation angle, the so-called full-dimension MIMO can house many more antennas with the same form factor [14]. As a side benefit, tailored vertical beams increase the signal power and reduce interference to users in neighboring cells.

4. *Channel models*: Together with the architectural issues, there are those related to channel models, which to be sound require extensive field measurements. Antenna correlations and couplings for massive arrays with relevant topologies must be determined, and a proper modeling of their impact must be established; in particular, the degree of actual channel orthogonalization in the face of such nonidealities must be verified. And,

for full-dimensional MIMO, besides azimuth, the modeling needs to incorporate elevation, which is a dimension on which far less data exists concerning power spectra and angle spreads [15], [16].

5. *Coexistence with small cells*: Massive MIMO BSs would most likely have to coexist with tiers of small cells, which would not be equipped with massive MIMO due to their smaller form factor. Although the simplest alternative is to segregate the corresponding transmissions in frequency, the large number of excess antennas at massive MIMO BSs may offer the opportunity of spatial nulling and interference avoidance with relative simplicity and little penalty. As networks become dense and more traffic is offloaded to small cells, the number of active users per cell will diminish and the need for massive MIMO may decrease. Aspects such as cost and backhaul will ultimately determine the balance between these complementary ideas [17], [18].

6. *Coexistence with mm-Wave*: Mm-Wave communication requires many antennas at the transmitter and at the receiver for beamsteering. The antennas are much smaller at these frequencies and thus they can conceivably fit very large numbers into portable devices. These antennas can indeed provide beamforming power gain but also MIMO opportunities. Any application of massive MIMO at mm-Wave frequencies would have to find the correct balance between power gain/interference reduction and parallelization [19], [20].

1.2 Thesis contributions

As mentioned above, this thesis focuses on the study of some applications of massive MIMO technology for future wireless networks. In particular massive MIMO is used in conjunction with two other current research topics for the future wireless networks: the mm-Wave frequencies and the distributed antenna systems.
A summary of the thesis contributions is provided in the following.

- **Chapter 2: Definition of massive MIMO Systems**
  The next chapter of the thesis gives a brief definition of massive MIMO systems and an historical and technical description of this technology. The conventional channel model and the uplink and downlink system model that is currently used in literature are introduced. The channel estimation in a massive MIMO network is described and the properties of channel hardening and favorable propagation are explained. Finally, an overview on the uplink and downlink spectral efficiency bounds for massive MIMO system is provided, using the tools proposed in the current literature.

- **Chapter 3: Massive MIMO at mm-Wave and at μ-Wave frequencies**
  The third chapter of the thesis presents an overview on the difference between massive MIMO at the conventional cellular frequencies and at mm-Wave. Initially conceived for conventional cellular frequencies in the sub-6 GHz range (μ-Wave), the massive MIMO concept has been then progressively extended to the case in which mm-Wave frequencies are used. However, due to different propagation mechanisms in urban scenarios, the resulting MIMO channel models at μ-Wave and mm-Wave are radically different. Six key basic differences are pinpointed in this chapter of the thesis, along with the implications that they have on the architecture and algorithms of the communication transceivers and on the attainable performance in terms of reliability and multiplexing capabilities.

  The material in this chapter is based on the journal article

  - S. Buzzi and C. D'Andrea "Massive MIMO 5G Cellular Networks: mm-Wave vs. μ-Wave Frequencies", *ZTE Communications-Special Topic: 5G New Radio (NR): Standard and Technology*, Vol. 15 No. S1, June 2017.

- **Chapter 4: Doubly massive MIMO mm-Wave systems**
  The fourth chapter of the thesis focuses on the "doubly massive MIMO" systems at mm-Wave, that is a system with a large number of antennas both at the transmitter and at the receiver. For complexity reasons and energy consumption issues, fully digital pre-coding and post-coding structures may turn out to be unfeasible, and thus suboptimal structures, making use of simplified hardware and a limited number of RF chains, have been investigated. This chapter considers and makes a comparative assessment of several suboptimal pre-coding and post-coding structures for a cellular MU-MIMO system with large number of antennas. Analytical formulas for the asymptotic achievable spectral efficiency (ASE) and for the global energy efficiency (GEE) of several beamforming structures are derived in the large number of antennas regime. Using the most recently available data for the energy consumption of phase shifters and switches, numerical results show that fully-digital beamformers may actually achieve a larger energy efficiency than lower-complexity solutions, as well as that low-complexity beam-steering purely analog beamforming may in some cases represent a good performance-complexity trade-off solution.

  The material in this chapter is based on the journal article

  - S. Buzzi and C. D'Andrea, "Energy Efficiency and Asymptotic Performance Evaluation of Beamforming Structures in Doubly Massive MIMO mmWave Systems", in *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 385-396, June 2018,

on its conference versions published as

- S. Buzzi and C. D'Andrea, "Doubly Massive mmWave MIMO Systems: Using Very Large Antenna Arrays at Both Transmitter and Receiver", in *Proc. of 2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, 2016, pp. 1-6,
- S. Buzzi and C. D'Andrea, "Are mmWave Low-Complexity Beamforming Structures Energy-Efficient? Analysis of the Downlink MU-MIMO", in *Proc. of 2016 IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, 2016, pp. 1-6,

and on the talk

- C. D'Andrea, "The doubly massive MIMO regime in mmWave communications", talk delivered at the *2016 Tyrrhenian International Workshop on Digital Communications (TIW16)*, Sept. 2016.

- **Chapter 5: User-centric approach to cell-free massive MIMO networks**
The fifth chapter of the thesis focuses on the recently introduced cell-free (CF) massive MIMO architecture, wherein a very large number of distributed access points (APs), connected to a central processing unit (CPU), simultaneously and jointly serve a much smaller number of MSs. In this context, the term AP is used in order to distinguish it from the BS, that is characterized by an higher complexity with respect to the single AP in a distributed setup. The chapter extends the CF approach to the case in which both the APs and the MSs are equipped with multiple antennas, proposing a beamfoming scheme that, relying on the channel hardening effect, does not require channel estimation at the MSs. It contrasts the originally proposed formulation of CF massive MIMO with a user-centric (UC) approach wherein each MS is served only by a limited number of APs. Exploiting the framework of successive lower-bound maximization, the chapter of the thesis also proposes and analyzes power allocation strategies aimed at either sum-rate maximization or minimum-rate maximization, both for the uplink and downlink. Additionally, this chapter proposes advanced signal processing techniques to improve the performance of the UC architecture by using local partial zero forcing (PZF) pre-coding on the downlink and successive interference cancellation (SIC) on the uplink. The proposed schemes can be implemented locally, i.e., the channel estimates and the beamformers at the APs are computed and exploited locally, with no need to exchange information with the CPU. Numerical results show that the UC approach, which requires smaller backhaul overhead and is more scalable than the CF deployment, also achieves generally better performance than the CF approach for the vast majority of the users, especially on the uplink. Regarding the advanced signal processing

techniques here proposed, the results show that the proposed transceiver algorithms provide remarkable performance improvements with respect to the simple CM pre-coding on the downlink and MR post-coding on the uplink.

The material in this chapter is based on the journal articles

- S. Buzzi and C. D'Andrea, "Cell-Free Massive MIMO: User-Centric Approach", in *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706-709, Dec. 2017,
- S. Buzzi, C. D'Andrea, A. Zappone and C. D'Elia "User-Centric 5G Cellular Networks: Resource Allocation and Comparison with the Cell-Free Massive MIMO Approach", submitted to *IEEE Transactions on Wireless Communications*, July 2018,

and on the conference papers

- S. Buzzi and C. D'Andrea, "User-Centric Communications versus Cell-free Massive MIMO for 5G Cellular Networks", in *Proc. of WSA 2017; 21th ITG International Workshop on Smart Antennas*, Berlin, Germany, Mar. 2017,
- S. Buzzi, C. D'Andrea and C. D'Elia, "User-Centric Cell-Free Massive MIMO with Interference Cancellation and Local ZF Downlink Precoding", in *Proc. of 2018 15th International Symposium on Wireless Communication Systems (ISWCS)*, Lisbon, Aug. 2018, pp. 1-5.

- **Chapter 6: Conclusions and future developments**
  The last chapter of the thesis gives some concluding remarks on the contribution of this work. The content of each chapter is briefly summarized and discussed. Finally, a brief overview on the future developments of the massive MIMO technology is given.

1.3 Other contributions

In addition to the articles listed above, the author has had other publications/submissions during the Ph.D. studies. In the following a brief abstract of each paper is given.

1.3.1 Journals

- S. Buzzi, C. D'Andrea, T. Foggi, A. Ugolini and G. Colavolpe, "Single-Carrier Modulation Versus OFDM for Millimeter-Wave Wireless MIMO", in *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1335-1348, March 2018.

  This paper presents results on the achievable spectral efficiency and on the energy efficiency for a wireless MIMO link operating at mm-Wave in a typical 5G

scenario. Two different single-carrier modem schemes are considered, i.e., a traditional modulation scheme with linear equalization at the receiver, and a single-carrier modulation with cyclic prefix, frequency-domain equalization and fast Fourier transform (FFT)-based processing at the receiver; these two schemes are compared with a conventional MIMO orthogonal frequency division multiplexing transceiver structure. The analysis reported in the paper jointly takes into account the peculiar characteristics of MIMO channels at mm-Wave frequencies, the use of hybrid (analog-digital) pre-coding and post-coding beamformers, the finite cardinality of the modulation structure, and the non-linear behavior of the transmitter power amplifiers. Results show that the best performance is achieved by single-carrier modulation with time-domain equalization, which exhibits the smallest loss due to the non-linear distortion, and whose performance can be further improved by using advanced equalization schemes. Results also confirm that performance gets severely degraded when the link length exceeds 90-100 m and the transmit power falls below 0 dBW.

- S. Buzzi, C. D'Andrea, Dejian Li and Shulan Feng "MIMO-UFMC Transceiver Schemes for Millimeter Wave Wireless Communications", submitted to *IEEE Transactions on Communications*, in second-round review, July 2018.

  The Universal Filtered MultiCarrier (UFMC) modulation is among the most considered solutions for the realization of beyond-Orthogonal Frequency Division Multiplexing (OFDM) air interfaces for future wireless networks. This paper focuses on the design and analysis of an UFMC transceiver equipped with multiple antennas and operating at mm-Wave carrier frequencies. The paper provides the full mathematical model of a MIMO-UFMC transceiver, taking into account the presence of hybrid analog/digital beamformers at both ends of the communication links. Then, several detection structures are proposed, both for the case of single-packet isolated transmission, and for the case of multiple-packet continuous transmission. In the latter situation, the paper also considers the case in which no guard time among adjacent packets is inserted, trading off an increased level of interference with higher values of spectral efficiency. At the analysis stage, the several considered detection structures and transmission schemes are compared in terms of bit-error-rate, root-mean-square-error, and system throughput. The numerical results show that the proposed transceiver algorithms are effective and that the linear MMSE data detector is capable of well managing the increased interference brought by the removal of guard times among consecutive packets, thus yielding throughput gains of about 10 - 13 %. The effect of phase noise at the receiver is also numerically assessed, and it is shown that the recursive implementation of the linear MMSE exhibits some degree of robustness against this disturbance.

1.3.2 Conferences

- S. Buzzi, C. D'Andrea, T. Foggi, A. Ugolini and G. Colavolpe, "Spectral Efficiency of MIMO Millimeter-Wave Links with Single-Carrier Modulation for 5G Networks", in *Proc. of WSA 2016; 20th International ITG Workshop on Smart Antennas*, Munich, Germany, 2016, pp. 1-8.

  Future wireless networks will extensively rely upon bandwidths centered on carrier frequencies larger than 10GHz. Indeed, recent research has shown that, despite the large pathloss, mm-Wave frequencies can be successfully exploited to transmit very large data-rates over short distances to slowly moving users. Due to hardware complexity and cost constraints, single-carrier modulation schemes, as opposed to the popular multi-carrier schemes, are being considered for use at mm-Wave frequencies. This paper presents preliminary studies on the achievable spectral efficiency on a wireless MIMO link operating at mm-Wave in a typical 5G scenario. Two different single-carrier modem schemes are considered, i.e., a traditional modulation scheme with linear equalization at the receiver, and a single-carrier modulation with cyclic prefix, frequency-domain equalization and FFT-based processing at the receiver. Results show that the former achieves a larger spectral efficiency than the latterand they also confirm that the spectral efficiency increases with the dimension of the antenna array, as well as that performance gets severely degraded when the link length exceeds 90-100 meters and the transmit power falls below 0dBW. Nonetheless, mm-Wave appear to be very suited for providing very large data-rates over short distances.

- S. Buzzi and C. D'Andrea, "Subspace Tracking Algorithms for Millimeter Wave MIMO Channel Estimation with Hybrid Beamforming", in *Proc. of WSA 2017; 21th International ITG Workshop on Smart Antennas*, Berlin, Germany, 2017, pp. 1-6.

  This paper proposes the use of subspace tracking algorithms for performing MIMO channel estimation at mm-Wave frequencies. Using a subspace approach, a protocol enabling the estimation of the right (resp. left) singular vectors at the transmitter (resp. receiver) side is developed; then, the projection approximation subspace tracking with deflation (PASTd) and the orthogonal Oja (OOJA) algorithms are adapted to the proposed framework and two channel estimation algorithms are obtained. The hybrid analog/digital nature of the beamformer is also explicitly taken into account at the algorithm design stage. Numerical results show that the proposed estimation algorithms are effective, and that they perform better than two relevant competing alternatives available in the open literature.

- S. Buzzi and C. D'Andrea, "Multiuser Millimeter Wave MIMO Channel Estimation with Hybrid Beamforming", poster presentation at *EuCNC 2017; European Conference on Networks and Communications*, Oulu, Finland, Jun. 2017.

This presentation focuses on MU-MIMO channel estimation and data transmission at mm-Wave frequencies. The proposed approach relies on the TDD protocol and is based on two distinct phases. First of all, the BS sends a suitable probing signal so that all the MSs, using a subspace tracking algorithm, can estimate the dominant left singular vectors of their BS-to-MS propagation channel. Then, each MS, using the estimated dominant left singular vectors as pre-coding beamformers, sends a suitable pilot sequence so that the BS can estimate the corresponding right dominant channel singular vectors and the corresponding eigenvalues. The low-complexity PASTd algorithm is used at the MSs for dominant subspace estimation, while pilot-matched (PM) and zero-forcing (ZF) reception is used at the BS. The proposed algorithms can be used in conjuction with an analog RF beamformer and are shown to exhibit very good performance.

- S. Buzzi, M. Lops, C. D'Andrea and C. D'Elia "Co-existence Between a Radar System and a Massive MIMO Wireless Cellular System", in *Proc. of 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Jun. 2018, pp. 1-5.

  In this paper the uplink of a massive MIMO communication system is considered, using 5G New Radio-compliant multiple access, which is to co-exist with a radar system using the same frequency band. The paper proposes a system model taking into account the reverberation (clutter) produced by the radar system at the massive MIMO receiver. Then, several linear receivers for uplink data-detection are proposed, ranging by the simple channel-matched beamformer to the zero-forcing and linear minimum mean square error receivers for clutter disturbance rejection. Results show that the clutter may have a strong effect on the performance of the cellular communication system, but the use of large-scale antenna arrays at the base station is key to provide increased robustness against it, at least as far as data-detection is concerned.

- C. D'Andrea, S. Buzzi, D. Li and S. Feng "Adaptive Data Detection in Phase-Noise Impaired MIMO-UFMC Systems at mmWave", in *Proc. of 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Bologna, Italy, Sep. 2018, pp. 231-235.

  This paper provides results on the use of the UFMC modulation scheme in MIMO wireless links operating at mm-Wave frequencies. After having provided mathematical details on the MIMO-UFMC signal model, taking into account both the hybrid analog/digital nature of the beamformers and the phase noise at the receiver, adaptive algorithms, based on the normalized least-mean-squares (NLMS) and on the recursive-least-squares (RLS) are proposed, for low-complexity data-detection. The proposed transceiver also includes a channel independent transmit beamformer, so as to avoid the need for channel state information at the transmitter, as well as a transmission format wherein no spacing between consecutive

data packets is adopted, in order to have increased spectral efficiency. Numerical results show that the proposed algorithms exhibit a certain amount of robustness to the phase noise and achieve satisfactory performance.

- S. Buzzi, C. D'Andrea, D. Li and S. Feng "MIMO-UFMC Transceiver Schemes for Millimeter Wave Wireless Communications", in *Proc. of Chinacom 2018 - 13th EAI International Conference on Communications and Networking in China*, October 2018.

  This paper provides results on the use of UFMC modulation scheme in MIMO wireless links operating at mm-Wave frequencies. First of all, full mathematical details on the processing needed to realize a MIMO-UFMC transceiver at mm-Wave, taking into account also the hybrid analog/digital nature of the beamformers, are given. Then, several reception structures are proposed, considering also the case of continuous packet transmission with no guard intervals among the packets. In particular, an adaptive low complexity MMSE receiver is proposed that is shown to achieve very satisfactory performance. A channel independent transmit beamformer is also considered, in order to avoid the need for channel state information at the transmitter. Numerical results show that the proposed transceiver schemes are effective, as well as that the continuous packet transmission scheme, despite increased interference, attains the highest values of system throughput.

Chapter 2

Definition of massive MIMO Systems

2.1 Introduction

This chapter of the thesis briefly defines many of the basic concepts related to massive MIMO technology [21]. A formal definition of massive MIMO networks is provided, the correlated and uncorrelated channel models are described, the uplink and downlink system models used in current literature are detailed. The channel estimation in a massive MIMO network is described and, finally, the properties of channel hardening and favorable propagation are explained. Finally, an overview on the uplink and downlink spectral efficiency bounds for massive MIMO system is provided, using the tools proposed in the current literature.
A highly spectrally efficient coverage tier in a cellular network can be characterized as follows:

- It uses space division multiple access (SDMA) to achieve a multiplexing gain by serving multiple MSs on the same time-frequency resource.
- It has more BS antennas than MSs per cell to achieve efficient interference suppression. If the number of MSs grows in a cell, the BS should be upgraded so that the number of antennas increases proportionally.
- It operates in time division duplex (TDD) mode to limit the channel state information (CSI) acquisition overhead, due to the multiple antennas, and to not rely on parametrizable channel models.

The massive MIMO technology embraces these design guidelines, making it an efficient way to achieve high spectral efficiency in the coverage tier of future wireless networks. A *concise definition of massive MIMO* can be the following [4], [21].

Definition: A *massive MIMO network* is a multicarrier cellular network with L cells that operate according to a synchronous TDD protocol[1]. The $j-$th BS is

[1] A synchronous TDD protocol refers to a protocol in which uplink and downlink transmissions within different cells are synchronized.

equipped with $M_j \gg 1$ antennas and it communicates with K_j single-antenna MSs simultaneously on each time/frequency sample, with antenna-MS ratio $\frac{M_j}{K_j} > 1$. Each BS operates individually and processes its signals using linear receive post-coding and linear transmit pre-coding.

The key point of the massive MIMO system is not which multicarrier modulation scheme is used, but that the frequency resources are divided into *flat-fading* subcarriers. The *coherence bandwidth* B_c describes the frequency interval over which the channel responses are approximately constant. One or multiple subcarriers fit(s) into the coherence bandwidth, thus the channel observed on adjacent subcarriers are either approximately equal or closely related through a deterministic transformation. Hence, there is generally no need to estimate the channel on every subcarrier. Similarly, the time variations of the channels are small between adjacent samples and the *coherence time* T_c describes the time interval over which the channel responses are approximately constant. A *coherence block* consists of a number of subcarriers and time samples over which the channel response can be approximated as constant and flat-fading. If the coherence bandwidth is B_c and the coherence time is T_c, then each coherence block contains $\tau_c = B_c T_c$ complex-valued samples.

2.2 Channel model

In the following of this chapter, following the current literature on massive MIMO system, single-antenna MSs are assumed and the cell index is explicitly reported, i.e., all the processing is made after the MS-BS selection. The channel response between the $k-$th single-antenna MS in $\ell-$th cell and the BS in $j-$th cell is denoted by $\mathbf{h}_{\ell,k,j} \in \mathbb{C}^{M_j}$, where each of the elements corresponds to the channel response from the MS to one of the BS's M_j antennas. The channel response is the same in both uplink and downlink of a coherence block. The uplink channel is denoted as $\mathbf{h}_{\ell,k,j}$ and the downlink channel as $\mathbf{h}_{\ell,k,j}^H$, although there is only a transpose and not any complex conjugate in practice. The additional conjugation does not change the spectral efficiency or any other performance metric, but simplifies the notation. Since the channel response is a vector, it is characterized by its norm and its direction in the vector space. Both are random variables in a fading channel. The channel model characterizes their respective distribution and statistical independence/dependence. A generic fading channel $\mathbf{h} \in \mathbb{C}^M$ is *spatially uncorrelated* if the channel gain $\|\mathbf{h}\|^2$ and the channel direction $\frac{\mathbf{h}}{\|\mathbf{h}\|}$ are independent random variables, and the channel direction is uniformly distributed over the unit-sphere in $\mathbb{C}^M$. The channel is otherwise *spatially correlated*. The vast majority of the massive MIMO literature concentrates on *correlated Rayleigh fading channels* such that

$$\mathbf{h}_{\ell,k,j} \sim \mathcal{CN}\left(\mathbf{0}_{M_j}, \mathbf{R}_{\ell,k,j}\right). \tag{2.1}$$

In the case of $\mathbf{R}_{\ell,k,j} = \beta_{\ell,k,j}\mathbf{I}_{M_j}$ the channel is spatially uncorrelated, where $\beta_{\ell,k,j}$ refers to as the large-scale fading coefficient of the channel between the $k-$th MS in the $\ell-$th cell and the $j-$th BS.

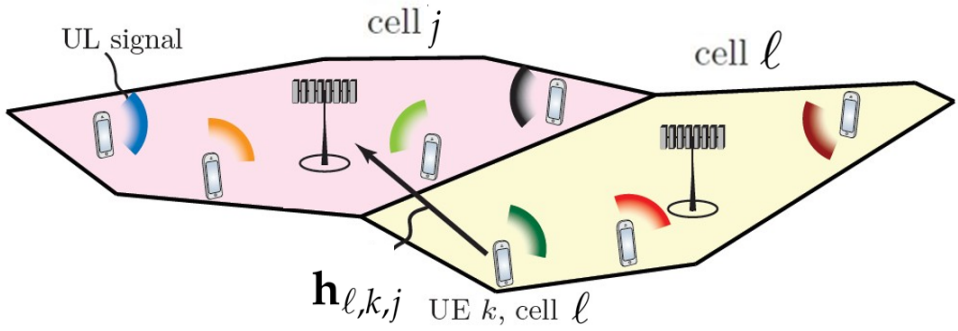

FIGURE 2.1: Illustration of the uplink massive MIMO transmission in the $j-$th cell and the $\ell-$th cell. The channel vector between the $j-$th BS and the $k-$th MS is called $\mathbf{h}_{\ell,k,j}$ [21].

2.3 System model

In this section, a generic system model for the uplink and downlink is provided.

- **Uplink data transmission**: The uplink transmission in massive MIMO is illustrated in Fig. 2.1. The received uplink signal $\mathbf{y}_j \in \mathbb{C}^{M_j}$ at the BS in the $j-$th cell is modeled as

$$\begin{aligned}\mathbf{y}_j &= \sum_{\ell=1}^{L}\sum_{k=1}^{K_\ell}\sqrt{p_{\ell,k}}\mathbf{h}_{\ell,k,j}s_{\ell,k} + \mathbf{w}_j \\ &= \underbrace{\sum_{k=1}^{K_j}\sqrt{p_{j,k}}\mathbf{h}_{j,k,j}s_{j,k}}_{\text{desired signal}} + \underbrace{\sum_{\substack{\ell=1\\ \ell\neq j}}^{L}\sum_{i=1}^{K_\ell}\sqrt{p_{\ell,i}}\mathbf{h}_{\ell,i,j}s_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{\mathbf{w}_j}_{\text{noise}},\end{aligned} \tag{2.2}$$

where $\mathbf{w}_j \sim \mathcal{CN}\left(\mathbf{0}_{M_j}, \sigma^2_{\text{UL}}\mathbf{I}_{M_j}\right)$ is independent additive receiver noise with zero mean and variance σ^2_{UL} and $p_{\ell,k}$ is the power transmitted from the $k-$th MS in the $\ell-$th cell. The uplink signal from the $k-$th MS in the $\ell-$th cell is denoted by $s_{\ell,k} \in \mathbb{C}$ and has unit power, i.e., $\mathbb{E}\left[|s_{\ell,k}|^2\right] = 1$, irrespective of whether it is a random payload data signal $s_{\ell,k} \sim \mathcal{CN}(0,1)$ or a deterministic pilot signal with $|s_{\ell,k}|^2 = 1$. The channels are constant within a coherence block, while the signals and noise take new realization at every sample. During data transmission, the BS in the $j-$th cell selects the receive post-coding vector denoted, at this point, as $\mathbf{d}_{j,k} \in \mathbb{C}^{M_j}$ to separate

the signal from its $k-$th desired MS from the interference as

$$\begin{aligned}\mathbf{d}_{j,k}^H \mathbf{y}_j &= \underbrace{\sqrt{p_{j,k}}\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j} s_{j,k}}_{\text{desired signal}} + \underbrace{\sum_{\substack{i=1 \\ i\neq k}}^{K_j} \sqrt{p_{j,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{j,i,j} s_{j,i}}_{\text{intra-cell interference}} \\ &+ \underbrace{\sum_{\substack{\ell=1 \\ \ell\neq j}}^{L} \sum_{i=1}^{K_\ell} \sqrt{p_{\ell,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{\ell,i,j} s_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{\mathbf{d}_{j,k}^H \mathbf{w}_j}_{\text{noise}} .\end{aligned} \tag{2.3}$$

The selection of post-coding vectors, based on estimated channels, and the corresponding uplink spectral efficiency bounds will be detailed in the following section. Note that receiver post-coding is a *linear processing* scheme that is also known as *linear detection*.

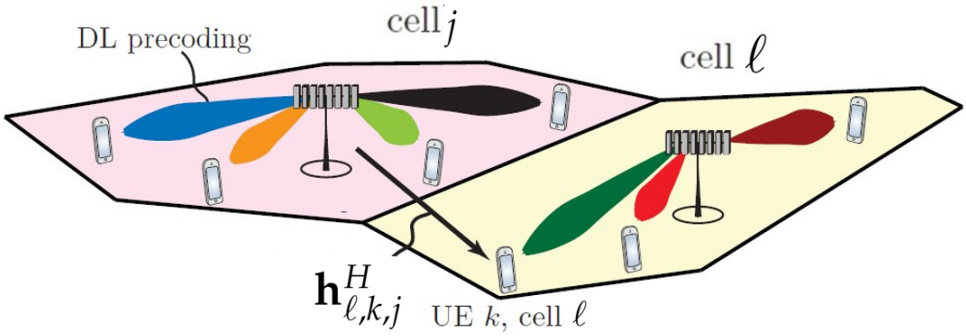

FIGURE 2.2: Illustration of the downlink massive MIMO transmission in the $j-$th cell and the $\ell-$th cell. The channel vector between the $j-$th BS and the $k-$th MS is called $\mathbf{h}_{\ell,k,j}$ [21].

- **Downlink data transmission**: The downlink transmission in massive MIMO is illustrated in Fig. 2.2. The BS in the $\ell-$th cell transmits the signal

$$\mathbf{x}_\ell = \sum_{i=1}^{K_\ell} \sqrt{\rho_{\ell,i}}\mathbf{q}_{\ell,i}\zeta_{\ell,i}, \tag{2.4}$$

where $\zeta_{\ell,k} \sim \mathcal{CN}(0,1)$ is the downlink data signal intended for the $k-$th MS in the $\ell-$th cell and $\rho_{\ell,k}$ is the signal power. This signal is assigned to a transmit pre-coding vector $\mathbf{q}_{\ell,k} \in \mathbb{C}^{M_\ell}$ that determines the spatial directivity of the transmission. Note that the term $\rho_{\ell,k}$ takes into account also the normalization at the transmitter in order to allow the unit power constraint of the effective precoder, i.e., $\rho_{\ell,k} = \dfrac{p_{\ell,k}^{\text{DL}}}{\mathbb{E}\left[\|\mathbf{q}_{\ell,k}\|^2\right]}$. The received

signal $y_{j,k} \in \mathbb{C}$ at the $k-$th MS in the $j-$th cell is modeled as

$$\begin{aligned} y_{j,k} &= \sum_{\ell=1}^{L} \mathbf{h}_{j,k,\ell}^{H} \mathbf{x}_l + w_{j,k} \\ &= \sum_{\ell=1}^{L} \sum_{i=1}^{K_\ell} \sqrt{\rho_{\ell,i}} \mathbf{h}_{j,k,\ell}^{H} \mathbf{q}_{\ell,i} \zeta_{\ell,i} + w_{j,k} \\ &= \underbrace{\sqrt{\rho_{j,k}} \mathbf{h}_{j,k,j}^{H} \mathbf{q}_{j,k} \zeta_{j,k}}_{\text{desired signal}} + \underbrace{\sum_{\substack{i=1 \\ i \neq k}}^{K_j} \sqrt{\rho_{j,i}} \mathbf{h}_{j,k,j}^{H} \mathbf{q}_{j,i} \zeta_{j,i}}_{\text{intra-cell interference}} \\ &+ \underbrace{\sum_{\substack{\ell=1 \\ \ell \neq j}}^{L} \sum_{i=1}^{K_\ell} \sqrt{\rho_{\ell,i}} \mathbf{h}_{j,k,\ell}^{H} \mathbf{q}_{\ell,i} \zeta_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{w_{j,k}}_{\text{noise}}, \end{aligned} \tag{2.5}$$

where $w_{j,k} \sim \mathcal{CN}(0, \sigma_{\text{DL}}^2)$ is independent additive receiver noise with variance σ_{DL}^2. The channels are constant within a coherence block, while the signals and noise take new realization at every sample. The selection of transmit pre-coding vectors and the corresponding downlink spectral efficiency bounds will be detailed in the following section.

2.4 Channel estimation

To make efficient use of the massive number of antennas, each BS needs to estimate the channel responses from the MSs that are active in the current coherence block. It is particularly important for the $j-$th BS to have estimates of the channels from the MSs in the $j-$th cell. Channel estimates from interfering MSs in other cells can also be useful to perform interference suppression during data transmission. Each MS transmits a pilot sequence that spans τ_p samples. The pilot sequence of the $k-$th MS in the $j-$th cell is denoted by $\mathbf{\Phi}_{j,k} \in \mathbb{C}^{\tau_p}$. It is assumed to have unit-power, so that $\left\| \mathbf{\Phi}_{j,k} \right\|^2 = 1$. The signal received at the BS in the $j-$th cell is denoted as $\mathbf{Y}_{\text{p},j} \in \mathbb{C}^{M_j \times \tau_p}$ and can be written as

$$\mathbf{Y}_{\text{p},j} = \underbrace{\sum_{k=1}^{K_j} \sqrt{p_{j,k}} \mathbf{h}_{j,k,j} \mathbf{\Phi}_{j,k}^{T}}_{\text{desired pilots}} + \underbrace{\sum_{\substack{\ell=1 \\ \ell \neq j}}^{L} \sum_{i=1}^{K_\ell} \sqrt{p_{\ell,i}} \mathbf{h}_{\ell,i,j} \mathbf{\Phi}_{\ell,i}^{T}}_{\text{inter-cell pilots}} + \underbrace{\mathbf{W}_{\text{p},j}}_{\text{noise}}, \tag{2.6}$$

where $\mathbf{W}_{\text{p},j} \in \mathbb{C}^{M_j \times \tau_p}$ is the independent additive receiver noise with indipendent and identical distributed (i.i.d.) elements distributed as $\mathcal{CN}(0, \sigma_{\text{UL}}^2)$. To estimate the channel of a particular MS, the BS needs to know which pilot sequence this MS has transmitted. This is why the pilots are deterministic sequences and the pilot assignment is typically made when the MS connects to the BS; for example, using a random access procedure. Suppose, for the sake

of argument, that the $j-$th BS wants to estimate the channel $\mathbf{h}_{\ell,i,j}$ from an arbitrary $i-$th MS in the $\ell-$ th cell. The BS can then multiply/correlate $\mathbf{Y}_{\mathrm{p},j}$ with the pilot sequence $\mathbf{\Phi}_{\ell,i}$ of this MS, leading to the processed received pilot signal $\mathbf{y}^{\mathrm{p}}_{j,\ell,i} \in \mathbb{C}^{M_j}$, given as

$$\mathbf{y}^{\mathrm{p}}_{j,\ell,i} = \mathbf{Y}_{\mathrm{p},j}\mathbf{\Phi}^*_{\ell,i} = \sum_{\ell'=1}^{L}\sum_{i'=1}^{K_{\ell'}} \sqrt{p_{\ell',i'}}\mathbf{h}_{\ell',i',j}\mathbf{\Phi}^T_{\ell',i'}\mathbf{\Phi}^*_{\ell,i} + \mathbf{W}_{\mathrm{p},j}\mathbf{\Phi}^*_{\ell,i}\,. \tag{2.7}$$

For the $k-$th MS in the BS's own cell Eq. (2.7) can be expressed as

$$\begin{aligned}\mathbf{y}^{\mathrm{p}}_{j,j,k} = \mathbf{Y}_{\mathrm{p},j}\mathbf{\Phi}^*_{j,k} \;&= \underbrace{\sqrt{p_{j,k}}\mathbf{h}_{j,k,j}\mathbf{\Phi}^T_{j,k}\mathbf{\Phi}^*_{j,k}}_{\text{desired pilot}} + \underbrace{\sum_{\substack{i=1\\ i\neq k}}^{K_j}\sqrt{p_{j,i}}\mathbf{h}_{j,i,j}\mathbf{\Phi}^T_{j,i}\mathbf{\Phi}^*_{j,k}}_{\text{intra-cell pilots}} \\ &+ \underbrace{\sum_{\substack{\ell=1\\ \ell\neq j}}^{L}\sum_{i=1}^{K_\ell}\sqrt{p_{\ell,i}}\mathbf{h}_{\ell,i,j}\mathbf{\Phi}^T_{\ell,i}\mathbf{\Phi}^*_{j,k}}_{\text{inter-cell pilots}} + \underbrace{\mathbf{W}_{\mathrm{p},j}\mathbf{\Phi}^*_{j,k}}_{\text{noise}}.\end{aligned} \tag{2.8}$$

The second and third terms in Eq. (2.8) represent interference and contain inner products of the form $\mathbf{\Phi}^T_{\ell,i}\mathbf{\Phi}^*_{j,k}$ between the pilot of the desired MS and the pilot of another $i-$th MS in the $\ell-$th cell. If the pilot sequences of two UEs are orthogonal (i.e.,$\mathbf{\Phi}^T_{\ell,i}\mathbf{\Phi}^*_{j,k} = 0$), then the corresponding interference term in Eq. (2.8) vanishes and does not affect the estimation. Ideally, all pilot sequences should be orthogonal, but since the pilots are τ_p-dimensional vectors, for a given τ_p, we can only find a set of at most τ_p mutually orthogonal sequences. The finite length of the coherence blocks imposes the constraint $\tau_p < \tau_c$ that makes it impossible to assign mutually orthogonal pilots to all MSs in practice. Since longer pilots come at the price of having fewer samples for data transmission, it is non-trivial to optimize the pilot length; however, a rule-of-thumb is that τ_p should always be smaller than $\tau_c/2$ [10]. A practical assumption is that the network utilizes a set of τ_p mutually orthogonal pilot sequences, so that it is possible to define the set

$$\mathcal{P}_{j,k} = \{(\ell,i) : \mathbf{\Phi}_{\ell,i} = \mathbf{\Phi}_{j,k},\ \ell = 1,\ldots,L,\ i = 1,\ldots,K_\ell\}, \tag{2.9}$$

of all MSs that utilize the same pilot sequence as $k-$th MS in the $j-$th cell: $(\ell,i) \in \mathcal{P}_{j,k}$ means that the $i-$th MS in the $\ell-$th cell uses the same pilot as $k-$th MS in the $j-$th cell. Using the notation in Eq. (2.9), Eq. (2.8) simplifies into

$$\mathbf{y}^{\mathrm{p}}_{j,j,k} = \underbrace{\sqrt{p_{j,k}}\mathbf{h}_{j,k,j}}_{\text{desired pilot}} + \underbrace{\sum_{(\ell,i)\in\mathcal{P}_{j,k}\setminus(j,k)}\sqrt{p_{\ell,i}}\mathbf{h}_{\ell,i,j}}_{\text{interfering pilots}} + \underbrace{\mathbf{W}_{\mathrm{p},j}\mathbf{\Phi}^*_{j,k}}_{\text{noise}}. \tag{2.10}$$

Note that $\mathbf{y}^{\mathrm{p}}_{j,j,k} = \mathbf{y}^{\mathrm{p}}_{j,\ell,i}$, $\forall(\ell,i) \in \mathcal{P}_{j,k}$ since these MSs use the same pilot. We also note that $\mathbf{W}_{\mathrm{p},j}\mathbf{\Phi}^{*}_{j,k} \sim \left(\mathbf{0}_{M_j}, \sigma^2_{\mathrm{UL}}\mathbf{I}_{M_j}\right)$ since the pilot sequences are deterministic and $\left\|\mathbf{\Phi}_{j,k}\right\|^2 = 1$. The processed received signal $\mathbf{y}^{\mathrm{p}}_{j,j,k}$ in Eq. (2.10) is a sufficient statistic for estimating $\mathbf{h}_{j,k,j}$ since there is no loss in useful information as compared to using the originally received signal $\mathbf{Y}_{\mathrm{p},j}$ [22]. One of the most common techniques for the channel estimation in the massive MIMO literature is the minimum mean square error (MMSE) channel estimation. Focusing on the case of spatially uncorrelated channel, i.e., $\mathbf{h}_{\ell,i,j} \sim \mathcal{CN}\left(\mathbf{0}_{M_j}, \beta_{\ell,i,j}\mathbf{I}_{M_j}\right)$, the MMSE estimate of this channel based on the observation in Eq. (2.10) can be written as [22]

$$\widehat{\mathbf{h}}_{\ell,i,j} = \frac{\sqrt{p_{\ell,i}}\beta_{\ell,i,j}}{\sum\limits_{(\ell',i')\in\mathcal{P}_{\ell,i}} p_{\ell',i'}\beta_{\ell',i',j} + \sigma^2_{\mathrm{UL}}}\mathbf{y}^{\mathrm{p}}_{j,\ell,i} = c_{\ell,i,j}\mathbf{y}^{\mathrm{p}}_{j,\ell,i}. \tag{2.11}$$

The estimation error $\widetilde{\mathbf{h}}_{\ell,i,j} = \mathbf{h}_{\ell,i,j} - \widehat{\mathbf{h}}_{\ell,i,j}$ has correlation matrix [21], [22]

$$\mathbf{C}_{\ell,i,j} = \mathbb{E}\left[\widetilde{\mathbf{h}}_{\ell,i,j}\widetilde{\mathbf{h}}^{H}_{\ell,i,j}\right] = \beta_{\ell,i,j}\left(1 - p_{\ell,i}\beta_{\ell,i,j}c_{\ell,i,j}\right)\mathbf{I}_{M_j}. \tag{2.12}$$

An important property of the MMSE estimation is that $\widetilde{\mathbf{h}}_{\ell,i,j}$ and $\widehat{\mathbf{h}}_{\ell,i,j}$ are independent random variables, distributed as follows [21], [22]:

$$\begin{matrix} \widehat{\mathbf{h}}_{\ell,i,j} \sim \mathcal{CN}\left(\mathbf{0}_{M_j}, \beta_{\ell,i,j}\mathbf{I}_{M_j} - \mathbf{C}_{\ell,i,j}\right) \\ \widetilde{\mathbf{h}}_{\ell,i,j} \sim \mathcal{CN}\left(\mathbf{0}_{M_j}, \mathbf{C}_{\ell,i,j}\right) \end{matrix}. \tag{2.13}$$

2.5 Channel hardening and favorable propagation

Two important properties of multiantenna channels are *channel hardening* and *favorable propagation*. These properties are briefly detailed in the following.

- **Channel hardening**: It makes a fading channel behave as deterministic. This property alleviates the need for combating small-scale fading (e.g., by adapting the transmit powers) and improves the downlink channel gain estimation. A propagation channel $\mathbf{h}_{\ell,k,j}$ provides asymptotic channel hardening if

$$\frac{\left\|\mathbf{h}_{\ell,k,j}\right\|^2}{\mathbb{E}\left[\left\|\mathbf{h}_{\ell,k,j}\right\|^2\right]} \to 1 \text{ almost surely as } M_j \to \infty. \tag{2.14}$$

  This definition says that the gain $\left\|\mathbf{h}_{\ell,k,j}\right\|^2$ of an arbitrary fading channel $\mathbf{h}_{\ell,k,j}$ is close to its mean value when there are many antennas.

- **Favorable propagation**: It makes the directions of two MS channels asymptotically orthogonal. This property makes it easier for the BS to mitigate interference between these MSs, which generally improves the spectral efficiency and makes it sufficient to use linear post-coding and pre-coding.

The pair of channels $\mathbf{h}_{\ell,i,j}$ and $\mathbf{h}_{j,k,j}$ to the $j-$th BS provide asymptotically favorable propagation if

$$\frac{\mathbf{h}_{\ell,i,j}^H \mathbf{h}_{j,k,j}}{\sqrt{\mathbb{E}\left[\left\|\mathbf{h}_{\ell,i,j}\right\|^2\right]\mathbb{E}\left[\left\|\mathbf{h}_{j,k,j}\right\|^2\right]}} \to 0 \text{ almost surely as } M_j \to \infty. \tag{2.15}$$

This definition says that the inner product of the normalized channels

$$\frac{\mathbf{h}_{\ell,i,j}}{\sqrt{\mathbb{E}\left[\left\|\mathbf{h}_{\ell,i,j}\right\|^2\right]}} \text{ and } \frac{\mathbf{h}_{j,k,j}}{\sqrt{\mathbb{E}\left[\left\|\mathbf{h}_{j,k,j}\right\|^2\right]}}$$

goes asymptotically to zero. Since the norms of the channels grow with M_j, favorable propagation does not imply that the inner product of $\mathbf{h}_{\ell,i,j}$ and $\mathbf{h}_{j,k,j}$ goes to zero; that is, the channel directions become orthogonal, but not the channel responses.

2.6 Spectral efficiency bounds for massive MIMO networks

In this section some spectral efficiency bounds for uplink and downlink are provided; these bounds have been proposed in the recent literature on the massive MIMO networks in order to evaluate their performance behaviour.
In order to clarify the need of spectral efficiency bounds evaluation, a simple exaplanation is provided in the following. Restricting the attention to linear beamforming (for downlink transmission), single data stream per MS, and independent channel coding of the MS data streams, a generic channel use of the underlying channel model is described by the Gaussian interference channel:

$$y_k = g_{k,k}s_k + \sum_{\substack{i=1 \\ i\neq k}}^{K} g_{k,i}s_i + w_k,\ k = 1,\ldots,K, \tag{2.16}$$

where y_k is the channel output observed by the $k-$th MS decoder, s_k is the coded information bearing symbol for the $k-$th MS (useful signal), $w_k \sim \mathcal{CN}(0,\sigma_w^2)$ is additive Gaussian noise, and $\{g_{k,i}\}$ are the *effective channel coefficients* resulting from the inner products of the transmit beamforming vectors with the MSs' channel vectors.
In general, when the coefficients $\{g_{k,i}\}$ are not perfectly known to the $k-$th MS receiver, for example because of imperfect CSI given by the channel estimation reported in Section 2.4, it is not clear what is "signal" and what is "interference" in Eq. (2.16). In particular, the intuitive notion of Signal-to-Interference plus

Noise Ratio (SINR), given by

$$\mathrm{SINR}_k = \frac{|g_{k,k}|^2}{\sum\limits_{\substack{i=1 \\ i\neq k}}^{K} |g_{k,i}|^2 + \sigma_w^2} \tag{2.17}$$

is in general not rigorously related to a corresponding notion of information theoretic achievable rate. Note that, Eq. (2.17) is rigorously related to a corresponding notion of information theoretic achievable rate in the case of perfect CSI.
In the following τ_d denotes the number of downlink data samples per coherence block, and consequently $\tau_u = \tau_c - \tau_p - \tau_d$ is the number of uplink data samples per coherence block.

2.6.1 Uplink spectral efficiency bounds

Focusing on the uplink spectral efficiency, in the following one upper bound and two lower bounds are derived.

- The upper bound for the uplink spectral efficiency is understood in the max-min sense, where the max is over the coding/decoding strategy of the $k-$th MS and the min is over all input distributions of the other MSs in the system. Considering Eq. (2.3), the spectral efficiency, measured in bit/s/Hz, is upper-bounded by [23]

$$\mathrm{SE}_{j,k,\mathrm{UL}}^{\mathrm{ub}} = \frac{\tau_u}{\tau_c}\mathbb{E}\left[\log_2\left(1 + \frac{p_{j,k}|\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}|^2}{\sum\limits_{\ell=1}^{L}\sum\limits_{\substack{i=1 \\ (i,\ell)\neq(j,k)}}^{K_\ell} p_{\ell,i}|\mathbf{d}_{j,k}^H \mathbf{h}_{\ell,i,j}|^2 + \sigma_{\mathrm{UL}}^2 \left\|\mathbf{d}_{j,k}\right\|^2}\right)\right], \tag{2.18}$$

  where the expectation is with respect to the channel realizations and the pre-log factor $\frac{\tau_d}{\tau_c}$ in Eq. (2.18) is the fraction of samples per coherence block that are used for uplink data. An important observation here and in the following is that in evaluating these bounds it is assumed that the decoder of the generic MS treats the multiuser interference as additional additive noise, i.e., the spectral efficiency is achieved by *Treating Interference as Noise* (TIN).

- The first spectral efficiency lower bound in the uplink is obtained rewriting Eq. (2.3) as

$$\mathbf{d}_{j,k}^H \mathbf{y}_j = \underbrace{\sqrt{p_{j,k}}\mathbf{d}_{j,k}^H \widehat{\mathbf{h}}_{j,k,j} s_{j,k}}_{\text{desired signal over estimated channel}} + \underbrace{\sqrt{p_{j,k}}\mathbf{d}_{j,k}^H \widetilde{\mathbf{h}}_{j,k,j} s_{j,k}}_{\text{desired signal over ``unknown'' channel}} + \underbrace{\sum_{\substack{i=1 \\ i\neq k}}^{K_j} \sqrt{p_{j,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{j,i,j} s_{j,i}}_{\text{intra-cell interference}} + \underbrace{\sum_{\substack{\ell=1 \\ \ell\neq j}}^{L} \sum_{i=1}^{K_\ell} \sqrt{p_{\ell,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{\ell,i,j} s_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{\mathbf{d}_{j,k}^H \mathbf{w}_j}_{\text{noise}}. \tag{2.19}$$

  The key difference in Eq. (2.19) is that the desired signal term has been divided into two parts: one that is received over the known estimated channel $\widehat{\mathbf{h}}_{j,k,j}$ from $k-$th MS in the $j-$th cell and one that is received over the unknown estimation error $\widetilde{\mathbf{h}}_{j,k,j}$ of the channel. The former part can be utilized straight away for signal detection, while the latter part is less useful since only the distribution of the estimation error is known. If MMSE channel estimation detailed in Section 2.4 is used, the uplink spectral efficiency of the $k-$th MS in the $j-$th cell is lower bounded by

$$\mathrm{SE}_{j,k,\mathrm{UL}}^{\mathrm{lb1}} = \frac{\tau_u}{\tau_c} \mathbb{E}\left[\log_2\left(1 + \mathrm{SINR}_{j,k,\mathrm{UL}}^{\mathrm{lb1}}\right)\right], \tag{2.20}$$

  with

$$\mathrm{SINR}_{j,k,\mathrm{UL}}^{\mathrm{lb1}} = \frac{p_{j,k}|\mathbf{d}_{j,k}^H \widehat{\mathbf{h}}_{j,k,j}|^2}{\sum_{\ell=1}^{L} \sum_{\substack{i=1 \\ (i,\ell)\neq(j,k)}}^{K_\ell} p_{\ell,i}|\mathbf{d}_{j,k}^H \widehat{\mathbf{h}}_{\ell,i,j}|^2 + \mathbf{d}_{j,k}^H \left(\sum_{\ell=1}^{L}\sum_{i=1}^{K_\ell} p_{\ell,i}\mathbf{C}_{\ell,i,j} + \sigma_{\mathrm{UL}}^2 \mathbf{I}_{M_j}\right)\mathbf{d}_{j,k}} \tag{2.21}$$

  and where the expectation is with respect to the channel estimates. The spectral efficiency in Eq. (2.20) holds under the assumption that the MMSE estimator is used for channel estimation and it can be computed by Monte Carlo simulations for any post-coding scheme, by generating many realizations of the instantaneous SINR in Eq. (2.21) [21].

- The second lower bound provided here is less tight but commonly used in research papers since it leads to closed-form expressions under the assumption of MMSE channel estimation and maximum ratio (MR) post-coding. The key idea behind this approach is to utilize the channel estimates only for computing the receive post-coding vectors, while this side information is not exploited in the signal detection. This simplification makes sense when there is substantial channel hardening. More precisely,

the receive combined signal in Eq. (2.3) is rewritten as

$$\mathbf{d}_{j,k}^H \mathbf{y}_j = \underbrace{\sqrt{p_{j,k}}\mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right] s_{j,k}}_{\text{desired signal over average channel}} + \underbrace{\sqrt{p_{j,k}}\left(\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j} - \mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right]\right) s_{j,k}}_{\text{desired signal over ``unknown'' channel}} + \underbrace{\sum_{\substack{i=1 \\ i\neq k}}^{K_j} \sqrt{p_{j,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{j,i,j} s_{j,i}}_{\text{intra-cell interference}} + \underbrace{\sum_{\substack{\ell=1 \\ \ell\neq j}}^{L} \sum_{i=1}^{K_\ell} \sqrt{p_{\ell,i}}\mathbf{d}_{j,k}^H \mathbf{h}_{\ell,i,j} s_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{\mathbf{d}_{j,k}^H \mathbf{w}_j}_{\text{noise}} . \tag{2.22}$$

by adding and subtracting $\sqrt{p_{j,k}}\mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right] s_{j,k}$. Only the part of the desired signal received over the average precoded channel $\mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right]$ is treated as the true desired signal. The part of $s_{j,k}$ received over the deviation from the mean value, $\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j} - \mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right]$, has zero mean and can thus be treated as an uncorrelated noise signal in the detection. The alternative capacity bound is referred to as the *use-and-then-forget* (UatF) bound since the channel estimates are used for post-coding and then effectively "forgotten" before the signal detection [24]. The uplink spectral efficiency of the $k-$th MS in the $j-$th cell is lower bounded by

$$\mathrm{SE}_{j,k,\mathrm{UL}}^{\mathrm{lb2}} = \frac{\tau_u}{\tau_c} \log_2\left(1 + \mathrm{SINR}_{j,k,\mathrm{UL}}^{\mathrm{lb2}}\right), \tag{2.23}$$

with

$$\mathrm{SINR}_{j,k,\mathrm{UL}}^{\mathrm{lb2}} = \frac{p_{j,k}\left|\mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right]\right|^2}{\sum_{\ell=1}^{L}\sum_{i=1}^{K_\ell} p_{\ell,i}\mathbb{E}\left[\left|\mathbf{d}_{j,k}^H \mathbf{h}_{\ell,i,j}\right|^2\right] - p_{j,k}\left|\mathbb{E}\left[\mathbf{d}_{j,k}^H \mathbf{h}_{j,k,j}\right]\right|^2 + \sigma_{\mathrm{UL}}^2 \mathbb{E}\left[\left\|\mathbf{d}_{j,k}\right\|^2\right]}, \tag{2.24}$$

where the expectations are with respect to the channel realizations. However, it does not require the use of MMSE channel estimation, but can be applied along with any channel estimator and any post-coding scheme. In fact, it can be applied with any channel distributions or even measured channels. Each of the expectations in Eq. (2.24) can be computed separately by means of Monte Carlo simulation. For MR post-coding, i.e., $\mathbf{d}_{j,k} = \widehat{\mathbf{h}}_{j,k,j}$, these expectations can be obtained in closed form as follows.

1) *Evaluation of* $\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{j,k,j}\right]$: Using the channel estimation procedure detailed in Section 2.4 we obtain

$$\begin{aligned}\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{j,k,j}\right] &= \mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \left(\widehat{\mathbf{h}}_{j,k,j} + \widetilde{\mathbf{h}}_{j,k,j}\right)\right] \\ &= \mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \widehat{\mathbf{h}}_{j,k,j}\right] = \sqrt{p_{j,k}}\beta_{j,k,j} c_{j,k,j} M_j .\end{aligned} \tag{2.25}$$

2) *Evaluation of* $\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{\ell,i,j}\right]$: Using Eqs. (2.10) and (2.11) we can write

$$\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{\ell,i,j}\right] = \mathbb{E}\left[c_{j,k,j}\left(\sum_{(\ell',i')\in\mathcal{P}_{j,k}} \sqrt{p_{\ell',i'}}\mathbf{h}_{\ell',i',j} + \widetilde{\mathbf{w}}_{j,k}\right)^H \mathbf{h}_{\ell,i,j}\right], \tag{2.26}$$

where $\widetilde{\mathbf{w}}_{j,k} = \mathbf{W}_{\mathrm{p},j}\mathbf{\Phi}_{j,k}^*$. Since the variance of a sum of independent random variables is equal to the sum of the variances, using the relation

$$\mathbb{E}\left[\left|\mathbf{h}_{\ell,i,j}^H \mathbf{h}_{\ell,i,j}\right|^2\right] = \beta_{\ell,i,j}^2 M_j(M_j+1), \tag{2.27}$$

and the definition in Eq. (2.11) we obtain

$$\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{\ell,i,j}\right] = p_{j,k} I_{(\ell,i)}^{(j,k)} M_j^2 \left(c_{j,k,j}\beta_{\ell,i,j}\right)^2 + \sqrt{p_{j,k}}c_{j,k,j}\beta_{\ell,i,j}\beta_{j,k,j}M_j, \tag{2.28}$$

with

$$I_{(\ell,i)}^{(j,k)} = \begin{cases} 1 & \text{if } (\ell,i)\in\mathcal{P}_{j,k} \\ 0 & \text{otherwise} \end{cases}. \tag{2.29}$$

3) *Evaluation of* $\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \widehat{\mathbf{h}}_{\ell,i,j}\right]$: Similarly to Eq. (2.25) we have

$$\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \widehat{\mathbf{h}}_{j,k,j}\right] = \sqrt{p_{j,k}}\beta_{j,k,j}c_{j,k,j}M_j. \tag{2.30}$$

Substituting Eqs. (2.25), (2.28) and (2.30) in Eq. (2.24) we obtain

$$\mathrm{SINR}_{j,k,\mathrm{UL}}^{\mathrm{lb2}} = \frac{\left(p_{j,k}\beta_{j,k,j}\right)^2 c_{j,k,j}M_j}{\underbrace{\sum_{\ell=1}^{L}\sum_{i=1}^{K_\ell}\sqrt{p_{j,k}}p_{\ell,i}\beta_{\ell,i,j}\beta_{j,k,j}}_{\text{non-coherent interference}} + \underbrace{\sum_{(\ell,i)\in\mathcal{P}_{j,k}\setminus(j,k)}\left(p_{\ell,i}\beta_{\ell,i,j}\right)^2 c_{j,k,j}M_j}_{\text{coherent interference}} + \sigma_{\mathrm{UL}}^2\sqrt{p_{j,k}}\beta_{j,k,j}}. \tag{2.31}$$

The closed-form spectral efficiency expression in Eq. (2.31) provides important insights into the basic behaviors of massive MIMO. The term at the numerator involves the transmit power multiplied with the term $c_{j,k,j}$, defined in Eq. (2.11). Hence, the estimation quality determines the signal strength and it is reduced by pilot contamination. At the denominator, the term referred to as *non-coherent interference* does not increase linearly with M_j, so that, when the number of antennas at the BS goes to infinity, it vanishes with respect to the term denoted as *coherent interference*. This term only involves the MSs in $\mathcal{P}_{j,k}\setminus(j,k)$, which are those using the same pilots as the desired MS and it is a consequence of the pilot contamination.

TABLE 2.1: Simulation parameters

Parameter	**Value**
Carrier frequency	1.9 GHz
Bandwidth	20 MHz
Noise figure (uplink and downlink)	9 dB
BS antenna height	15 m
MS antenna height	1.65 m
MS horizontal distribution	uniform
Channel estimation	MMSE detailed in Section 2.4
Thermal noise	-174 dBm/Hz spectral density
Power control	DL: no power control
	UL: no power control

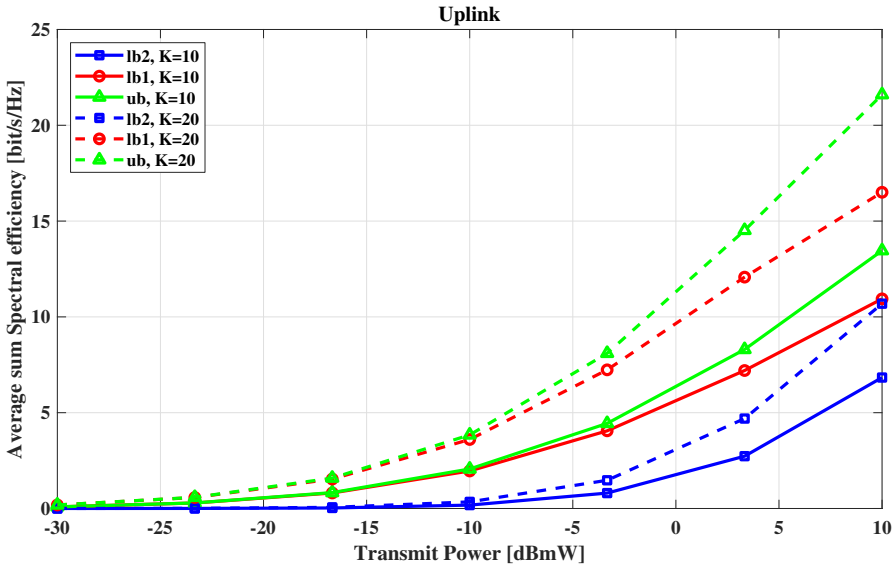

FIGURE 2.3: Average sum spectral efficiency in uplink versus the transmit power at the MSs of a system with $L = 4$ cells, $M_j = 100 \; \forall j = 1, \ldots, L$: comparison between the three bounds.

In Fig. 2.3 the average sum spectral efficiencies, measured in bit/s/Hz, are reported, using the three bounds for the uplink spectral efficiency detailed in this section, for two different values of number of MSs per cell. In this figure, a system with 4 cells is considered, i.e., $L = 4$, the number of antennas at the BS is the same for all the cells, i.e., $M_j = 100 \; \forall j = 1, \ldots, L$ and the number of MSs per cell is also the same for all the cells, i.e., $K_j = K \; \forall j = 1, \ldots, L$. The simulation parameters are reported in Table 2.1. In the figure the bound in Eq. (2.18) is denoted as "ub", the bound in Eq. (2.20) as "lb1" and the bound in Eq. (2.23) as "lb2". Inspectin the figure it is possible to note that the lower bound "lb1" is less tight than the bound "lb2", since the channel estimates are not utilized in the signal detection. The bound "lb2" yields a significantly better behavior for a

larger range of transmit power, however, for large values of transmit power the self-interference term becomes relevant and the curve flattens out and separates from the "ub". Increasing the number of MSs in the uplink it is possible to note that the performance of the system grows, so that the availability of multiple antennas at the BS can reduce the interference for the two values of number of MSs here considered.

2.6.2 Downlink spectral efficiency bounds

Focusing on the downlink spectral efficiency, in the following one upper bound and two lower bounds are explained.

- As in the case of uplink, the upper bound for the downlink spectral efficiency is understood in the max-min sense, where the max is over the coding/decoding strategy of the $k-$th MS and the min is over all input distributions of the other MSs in the system. Considering Eq. (2.5), the spectral efficiency is upper-bounded by [23]

$$\mathrm{SE}_{j,k,\mathrm{DL}}^{\mathrm{ub}} = \frac{\tau_d}{\tau_c}\mathbb{E}\left[\log_2\left(1+\frac{\rho_{j,k}\left|\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,k}\right|^2}{\sum\limits_{\ell=1}^{L}\sum\limits_{\substack{i=1\\(i,\ell)\neq(j,k)}}^{K_\ell}\rho_{\ell,i}\left|\mathbf{h}_{j,k,\ell}^H\mathbf{q}_{\ell,i}\right|^2+\sigma_{\mathrm{DL}}^2}\right)\right], \quad (2.32)$$

  where the expectation is with respect to the channel realizations and the pre-log factor $\frac{\tau_d}{\tau_c}$ in Eq. (2.32) is the fraction of samples per coherence block that are used for downlink data.

- The first spectral efficiency lower bound in the downlink is obtained rewriting Eq. (2.5) as

$$\begin{aligned} y_{j,k} \;=\; & \underbrace{\sqrt{\rho_{j,k}}\mathbb{E}\left[\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,k}\right]\zeta_{j,k}}_{\text{desired signal over average channel}} + \underbrace{\sqrt{\rho_{j,k}}\left(\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,k}-\mathbb{E}\left[\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,k}\right]\right)\zeta_{j,k}}_{\text{desired signal over ``unknown'' channel}} \\ & + \underbrace{\sum_{\substack{i=1\\i\neq k}}^{K_j}\sqrt{\rho_{j,i}}\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,i}\zeta_{j,i}}_{\text{intra-cell interference}} + \underbrace{\sum_{\substack{\ell=1\\\ell\neq j}}^{L}\sum_{i=1}^{K_\ell}\sqrt{\rho_{\ell,i}}\mathbf{h}_{j,k,\ell}^H\mathbf{q}_{\ell,i}\zeta_{\ell,i}}_{\text{inter-cell interference}} + \underbrace{w_{j,k}}_{\text{noise}}. \end{aligned} \quad (2.33)$$

  The first term in Eq. (2.33) is the desired signal received over the deterministic average precoded channel $\mathbb{E}\left[\mathbf{h}_{j,k,j}^H\mathbf{q}_{j,k}\right]$, while the remaining terms are random variables with realizations that are unknown to the MS. An achievable spectral efficiency can be computed by treating these terms as noise in the signal detection. This bound thus obtained is known as the *hardening bound* and it holds for any choice of pre-coding vectors and channel estimation schemes [21]. The downlink spectral efficiency of the $k-$th

MS in the $j-$th cell is lower bounded by

$$\mathrm{SE}_{j,k,\mathrm{DL}}^{\mathrm{lb1}} = \frac{\tau_d}{\tau_c} \log_2 \left(1 + \mathrm{SINR}_{j,k,\mathrm{DL}}^{\mathrm{lb1}}\right), \tag{2.34}$$

with

$$\mathrm{SINR}_{j,k,\mathrm{DL}}^{\mathrm{lb1}} = \frac{\rho_{j,k} \left|\mathbb{E}\left[\mathbf{q}_{j,k}^H \mathbf{h}_{j,k,j}\right]\right|^2}{\sum_{\ell=1}^{L}\sum_{i=1}^{K_\ell} \rho_{\ell,i} \mathbb{E}\left[\left|\mathbf{q}_{\ell,i}^H \mathbf{h}_{j,k,\ell}\right|^2\right] - \rho_{j,k}\left|\mathbb{E}\left[\mathbf{q}_{j,k}^H \mathbf{h}_{j,k,j}\right]\right|^2 + \sigma_{\mathrm{DL}}^2}. \tag{2.35}$$

The term in Eq. (2.35) is a deterministic scalar and the expression contains several expectations over the random channel realizations that can be computed separately by means of Monte Carlo simulation. If channel matched (CM) pre-coding with $\mathbf{q}_{j,k} = \widehat{\mathbf{h}}_{j,k,j}$ is used, based on the MMSE estimator detailed in Section 2.4, these expectations can be computed in closed form as follows.

1) *Evaluation of* $\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{j,k,j}\right]$: Using the channel estimation procedure detailed in Section 2.4 we obtain

$$\begin{aligned}\mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \mathbf{h}_{j,k,j}\right] &= \mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \left(\widehat{\mathbf{h}}_{j,k,j} + \widetilde{\mathbf{h}}_{j,k,j}\right)\right] \\ &= \mathbb{E}\left[\widehat{\mathbf{h}}_{j,k,j}^H \widehat{\mathbf{h}}_{j,k,j}\right] = \sqrt{p_{j,k}}\beta_{j,k,j} c_{j,k,j} M_j\,.\end{aligned} \tag{2.36}$$

2) *Evaluation of* $\mathbb{E}\left[\widehat{\mathbf{h}}_{\ell,i,\ell}^H \mathbf{h}_{j,k,\ell}\right]$: Using Eqs. (2.10) and (2.11) we can write

$$\mathbb{E}\left[\widehat{\mathbf{h}}_{\ell,i,\ell}^H \mathbf{h}_{j,k,\ell}\right] = \mathbb{E}\left[c_{\ell,i,\ell}\left(\sum_{(\ell',i')\in\mathcal{P}_{\ell,i}} \sqrt{p_{\ell',i'}}\mathbf{h}_{\ell',i',\ell} + \widetilde{\mathbf{w}}_{\ell,k}\right)^H \mathbf{h}_{j,k,\ell}\right]. \tag{2.37}$$

Since the variance of a sum of independent random variables is equal to the sum of the variances, using the relation

$$\mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \mathbf{h}_{j,k,\ell}\right|^2\right] = \beta_{j,k,\ell}^2 M_\ell (M_\ell + 1)\,, \tag{2.38}$$

and the definition in Eq. (2.11) we obtain

$$\mathbb{E}\left[\widehat{\mathbf{h}}_{\ell,i,\ell}^H \mathbf{h}_{j,k}^{\ell}\right] = p_{j,k} I_{(\ell,i)}^{(j,k)} M_\ell^2 \left(c_{\ell,i,\ell}\beta_{j,k,\ell}\right)^2 + \sqrt{p_{\ell,i}} c_{\ell,i,\ell}\beta_{\ell,i,\ell}\beta_{j,k,\ell} M_\ell\,, \tag{2.39}$$

where $I_{(\ell,i)}^{(j,k)}$ is defined as in Eq. 2.29. Substituting Eqs. (2.36) and (2.39) in Eq. (2.35) we obtain

$$\mathrm{SINR}_{j,k,\mathrm{DL}}^{\mathrm{lb1}} = \frac{\rho_{j,k} p_{j,k} \beta_{j,k,j}^2 c_{j,k,j} M_j^2}{\underbrace{\sum_{\ell=1}^{L} \sum_{i=1}^{K_\ell} \rho_{\ell,i} \sqrt{p_{\ell,i}} \beta_{j,k,\ell} c_{\ell,i,\ell} \beta_{\ell,i,\ell} M_\ell}_{\text{non-coherent interference}} + \underbrace{\sum_{(\ell,i) \in \mathcal{P}_{j,k} \setminus (j,k)} \rho_{\ell,i} p_{j,k} \beta_{j,k,\ell} \beta_{\ell,i,\ell} c_{\ell,i,\ell} M_\ell^2}_{\text{coherent interference}} + \sigma_{\mathrm{DL}}^2}. \tag{2.40}$$

The term at the numerator in Eq. (2.40) increases quadratically with M_j, this is the array gain from the pre-coding and his term involves $c_{j,k,j}$ that is related to the estimation quality. The first term in the numerator scales linearly with M_ℓ, so that it vanishes with respect to the second term when the number of antennas at the BSs grows. The second term in the denominator is the additional *coherent interference* that scales quadratically with M_ℓ and originates from the signals to MSs that share the same pilot; that is, pilot contamination also affects the downlink. In this case, the BS uses precoding to direct the signals towards the intended receivers, but partially also direct them towards the MSs that interfered with the pilot transmission.

- The second lower bound for the spectral efficiency in uplink is given by [23]

$$\mathrm{SE}_{j,k,\mathrm{DL}}^{\mathrm{lb2}} = \mathrm{SE}_{j,k,\mathrm{DL}}^{\mathrm{ub}} - \frac{\tau_d}{\tau_c^2} \sum_{\ell=1}^{L} \sum_{i=1}^{K_\ell} \log_2 \left(1 + \frac{\tau_c}{\sigma_{\mathrm{DL}}^2} \rho_{\ell,i} \mathrm{Var} \left[\mathbf{h}_{j,k,\ell}^H \mathbf{q}_{\ell,i} \right] \right). \tag{2.41}$$

The expectations over the random channel realizations in Eq. (2.41) can be computed separately by means of Monte Carlo simulation. If CM precoding with $\mathbf{q}_{j,k} = \widehat{\mathbf{h}}_{j,k,j}$ is used, based on the MMSE estimator detailed in Section 2.4, the variance in the second term can be computed as follows:

$$\begin{aligned} \mathrm{Var} \left[\mathbf{h}_{j,k,\ell}^H \widehat{\mathbf{h}}_{\ell,i,\ell} \right] &\overset{(a)}{=} \mathbb{E} \left[\left| \mathbf{h}_{j,k,\ell}^H \widehat{\mathbf{h}}_{\ell,i,\ell} \right|^2 \right] \\ &\overset{(b)}{=} \mathbb{E} \left[\left| \mathbf{h}_{j,k,\ell}^H c_{\ell,i,\ell} \left(\sqrt{p_{\ell,i}} \mathbf{h}_{\ell,i,\ell} + \sum_{(\ell',i') \in \mathcal{P}_{\ell,i} \setminus (\ell,i)} \sqrt{p_{\ell',i'}} \mathbf{h}_{\ell',i',\ell} + \widetilde{\mathbf{w}}_{\ell,i} \right) \right|^2 \right], \end{aligned} \tag{2.42}$$

where (a) follows from the fact that both the channel and its MMSE estimate are zero mean, and (b) is obtained substituting Eqs. (2.10) and (2.11) in Eq. (2.42). Now, the two cases $(j,k) \in \mathcal{P}_{\ell,i}$ and $(j,k) \notin \mathcal{P}_{\ell,i}$ are analized separately. Starting from the first case and using the fact that the variance

of a sum of independent random variables is equal to the sum of the variances Eq. (2.42) is written as

$$\begin{aligned}
&\mathrm{Var}\left[\mathbf{h}_{j,k,\ell}^H \widehat{\mathbf{h}}_{\ell,i}^{\ell}\right] = c_{\ell,i,\ell}^2 p_{j,k} \mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \mathbf{h}_{j,k,\ell}\right|^2\right] \\
&+c_{\ell,i,\ell}^2 \sum_{(\ell',i')\in\mathcal{P}_{\ell,i}\setminus(j,k)} p_{\ell',i'} \mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \mathbf{h}_{\ell',i',\ell}\right|^2\right] + c_{\ell,i,\ell}^2 \mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \widetilde{\mathbf{w}}_{\ell,i}\right|^2\right] \\
&\overset{(c)}{=} c_{\ell,i,\ell}^2 p_{j,k} \beta_{j,k,\ell}^2 M_\ell (M_\ell + 1) + \\
&+c_{\ell,i,\ell}^2 \sum_{(\ell',i')\in\mathcal{P}_{\ell,i}\setminus(j,k)} p_{\ell',i'} \beta_{j,k,\ell} \beta_{\ell',i',\ell} M_\ell + \sigma_{\mathrm{DL}}^2 c_{\ell,i,\ell}^2 \beta_{j,k,\ell} M_\ell \\
&\overset{(d)}{=} c_{\ell,i,\ell}^2 p_{j,k} \beta_{j,k,\ell}^2 M_\ell^2 + c_{\ell,i,\ell} \beta_{j,k,\ell} \beta_{\ell,i,\ell} \sqrt{p_{\ell,i}} M_\ell \,,
\end{aligned} \tag{2.43}$$

where (c) follows from the independence between different channels and channel and noise and from Eq. 2.38. The equality (d) follows from definition of $c_{\ell,i,\ell}$ in Eq. (2.11).

Considering now the case $(j,k) \notin \mathcal{P}_{\ell,i}$ Eq. (2.42) is written, using similar arguments as in Eq. (2.43), as

$$\begin{aligned}
&\mathrm{Var}\left[\mathbf{h}_{j,k,\ell}^H \widehat{\mathbf{h}}_{\ell,i,\ell}\right] = c_{\ell,i,\ell}^2 \sum_{(\ell',i')\in\mathcal{P}_{\ell,i)}} p_{\ell',i'} \mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \mathbf{h}_{\ell',i',\ell}\right|^2\right] \\
&+c_{\ell,i,\ell}^2 \mathbb{E}\left[\left|\mathbf{h}_{j,k,\ell}^H \widetilde{\mathbf{w}}_{\ell,i}\right|^2\right] \\
&= c_{\ell,i,\ell}^2 \sum_{(\ell',i')\in\mathcal{P}_{\ell,i}} p_{\ell',i'} \beta_{j,k,\ell} \beta_{\ell',i',\ell} M_\ell + \sigma_{\mathrm{DL}}^2 c_{\ell,i,\ell}^2 \beta_{j,k,\ell} M_\ell \\
&= c_{\ell,i,\ell} \beta_{j,k,\ell} \beta_{\ell,i,\ell} \sqrt{p_{\ell,i}} M_\ell \,.
\end{aligned} \tag{2.44}$$

Using Eqs. (2.43) and (2.44), Eq. (2.42) can be rewritten as

$$\mathrm{Var}\left[\mathbf{h}_{j,k,\ell}^H \widehat{\mathbf{h}}_{\ell,i,\ell}\right] = c_{\ell,i,\ell} \beta_{j,k,\ell} \beta_{\ell,i,\ell} \sqrt{p_{\ell,i}} M_\ell + I_{(j,k)}^{(\ell,i)} c_{\ell,i,\ell}^2 p_{j,k} \beta_{j,k,\ell}^2 M_\ell^2 \,, \tag{2.45}$$

where $I_{(j,k)}^{(\ell,i)}$ is defined as in Eq. (2.29).

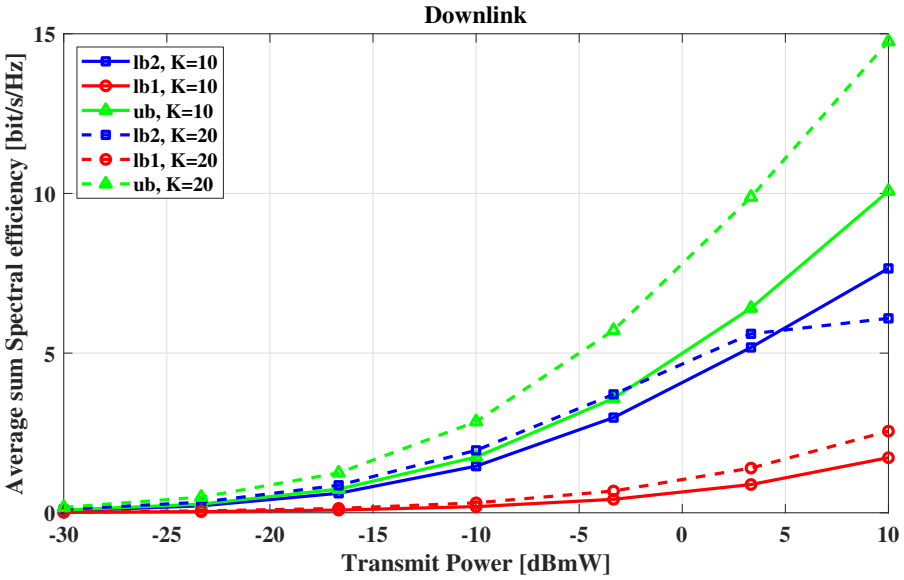

FIGURE 2.4: Average sum spectral efficiency in downlink versus the transmit power at the MSs of a system with $L = 4$ cells, $M_j = 100 \; \forall \, j = 1, \ldots, L$: comparison between the three bounds.

In Fig. 2.4 the average sum spectral efficiencies, measured in bit/s/Hz, are reported, using the three bounds for the downlink spectral efficiency detailed in this section, for two different values of number of MSs per cell. In this figure, a system with 4 cells is considered, i.e., $L = 4$, the number of antennas at the BS is the same for all the cells, i.e., $M_j = 100 \; \forall \, j = 1, \ldots, L$ and the number of MSs per cell is also the same for all the cells, i.e., $K_j = K \; \forall \, j = 1, \ldots, L$. The simulation parameters are reported in Table 2.1. In the figure the bound in Eq. (2.18) is denoted as "ub", the bound in Eq. (2.20) as "lb1" and the bound in Eq. (2.23) as "lb2". Inspectin the figure it is possible to note that, as in the uplink, the lower bound "lb1" is less tight than the bound "lb2", since the channel estimates are not utilized in the signal detection. The bound "lb2" yields a significantly better behavior for a larger range of transmit power, expecially for $K = 10$, however, for high values of transmit power the self-interference term become relevant and the curve flattens out and separates from the "ub".

Chapter 3

Massive MIMO at mm-Wave and at μ-Wave frequencies

Referred article published in the *ZTE Communications-Special Topic: 5G New Radio (NR): Standard and Technology* 2017.

3.1 Introduction

Future wireless networks are expected to provide 1000x improvement on the supported data rate, as compared to current LTE networks. Enhanced mobile broadband (eMBB) is one of the key use-cases for the development of the new standard 5G New Radio for the next generation of mobile wireless networks. As explained in Chapter 1, three are the main factors that will permit implementing eMBB services and realizing the Gbit/s mobile wireless experience[1]: (a) the reduction in the size of the radio-cells, so that a larger data-rate density can be achieved; (b) the use of large-scale antenna arrays at the BSs, i.e., MIMO [25], so that several users can be multiplexed in the same time-frequency resource slot through MU-MIMO techniques; and (c) the use of carrier frequencies in the range 10–100 GHz, a.k.a. mm-Waves[2] [26], so that larger bandwidths become available. The factor (a), i.e., the densification of the network, is actually a trend that can be observed for some decades, in the sense that the size of the radio cells has been progressively reduced over time from one generation of cellular networks to the next one. Differently, factor (b) can be seen as a sort of 4.5G technology, in the sense that the latest Third Generation Partnership Project (3GPP) LTE releases already include the possibility to equip BS with antenna arrays of up to 64 elements. This trend will certainly continue in the 5G New Radio (5G-NR) standard, since the potentialities of massive MIMO are currently being tested worldwide in a number of real-world experiments (for instance, [27] and [28]). The use of mm-Waves, on the contrary, is a more recent technology, at least as far as wireless cellular applications are concerned, but mm-Waves can be certainly classified as a technology of the next generations of cellular networks.

[2]Even though mm-Waves is a term that historically refers to the range 30-300 GHz, in the recent literature about future wireless networks the term is used to refer to frequency above-6 GHz, in contrapposition to the usual cellular frequencies located below 6 GHz.

Focusing on the massive MIMO technology, most of the research and experimental work has mainly considered its use at conventional cellular frequencies (e.g. sub-6 GHz). Such a range of frequencies is denoted here as μ-Wave, to contrast them with the above-6 GHz frequencies that it is denoted as mm-Wave. Only recently, the combination of the massive MIMO concept with the use of mm-Wave frequency bands has started being considered [19], [29]. As a matter of fact, the channel propagation mechanisms at μ-Wave frequencies are completely different from those at mm-Waves. The propagation conditions are less favorable in mm-Wave bands, making the beamforming gain offered by massive MIMO an inescapable feature of such systems. There are, however, fundamental differences between how massive MIMO technology can be designed, implemented, and exploited in μ-Wave and mm-Wave bands [30]. For instance, the rich-scattering environment at μ-Wave in urban environments is observed [31], thus implying that the MIMO channel is customarily modeled as the product of a scalar constant when the shadowing effects and path loss times a matrix with i.i.d. entries are taken into account. At mm-Waves, instead, propagation is mainly based on Line-of-Sight (LOS) propagation and on one-hop reflections, and blockage phenomena are more frequent. To capture these mechanisms, a finite-rank clustered channel model is usually employed [32]–[34]. This chapter of the thesis compares massive MIMO systems at μ-Waves with massive MIMO systems at mm-Waves. It is observed that these two different channel models have key implications on the achievable performance, on the multiplexing capabilities of the channels themselves, on the beamforming strategies that can be employed, on the transceiver algorithms and on the adopted channel estimation procedures. Six key differences between massive MIMO systems at μ-Waves and massive MIMO systems at mm-Waves are thus identified and critically discussed.

The remainder of this chapter is organized as follows. Section 3.2 describes the considered transceiver model and the massive MIMO channel models at μ-Waves and at mm-Wave frequencies. Sections 3.3-3.8 describe six key differences between the massive MIMO channels at μ-Wave and at mm-Wave frequencies; numerical results are also shown here in order to provide experimental evidence of the theoretical discussion.

3.2 System and channel model

An understanding of the electromagnetic propagation is crucial when considering massive MIMO systems and frequencies up to mm-Wave bands. The channels behave fundamentally different from what it is used to in cellular networks, which exposes weaknesses in the channel modeling simplifications conventionally made. This section briefly illustrates the considered transceiver architecture and reviews the main characteristics of the MIMO wireless channel at μ-Wave

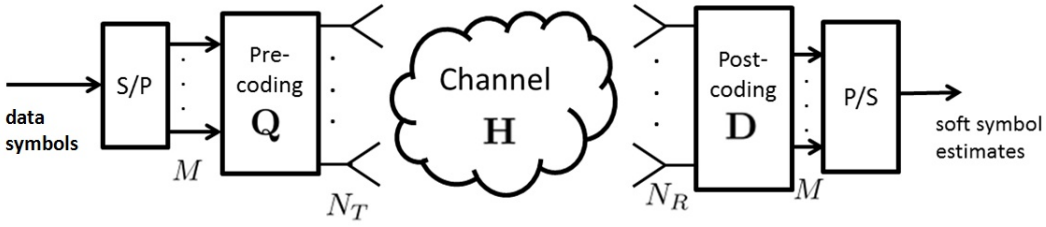

FIGURE 3.1: The considered transceiver model.

and mm-Wave carrier frequencies. In this chapter of the thesis it is considered a MIMO wireless link with N_T antennas at the transmitter and N_R antennas at the receiver. The distance between the transmitter and receiver is denoted by d, whereas the number of transmitted parallel data streams (i.e., the multiplexing order) is denoted by M. The considered transceiver model is shown in Fig. 3.1.

3.2.1 μ-Wave channel model

Radio channels below at μ-Waves have been widely studied for single-antenna and small-scale MIMO systems. The propagation depends on path-loss and shadowing, called large-scale fading, and multi-path propagation, resulting in small-scale fading. Assuming frequency-flat fading (i.e., either multipath may be neglected or it is nulled through the use of OFDM modulation), at μ-Waves the propagation channel is customarily modelled through an $(N_R \times N_T)-$dimensional matrix, whose $(i,j)-$th entry, $\left(\mathbf{H}_\mu\right)_{(i,j)}$, has the following structure[4]

$$\left(\mathbf{H}_\mu\right)_{(i,j)} = \sqrt{\beta} g_{i,j}\,, \tag{3.1}$$

where $g_{i,j}$ represents the small-scale (fast) fading between the $i-$th receive antenna and the $j-$th transmit antenna, and β represents the (slow) large-scale fading (shadowing) and the path-loss between the transmitter and the receiver. In a rich scattering environment, the coefficients $g_{i,j}$, $i = 1,\ldots,N_R$, $j = 1,\ldots,N_T$ are i.i.d. $\mathcal{CN}(0,1)$ random variables. The factor β in Eq. (3.1) is assumed constant across the transmit and receive antennas (i.e., it does not depend on the indices i, j), and is usually expressed as:

$$\beta = 10^{\frac{\mathrm{PL}}{10}} 10^{\frac{\sigma_{\mathrm{sh}} z}{10}}, \tag{3.2}$$

where PL represents the path loss (expressed in dB) and $10^{\frac{\sigma_{\mathrm{sh}} z}{10}}$ represents the shadow fading with standard deviation σ_{sh}, while $z \sim \mathcal{N}(0,1)$. With regard to the path loss PL, several models have been derived over the years, based on theoretical models and/or on empirical heuristics. According to the popular

three-slope model [35], [36], the path loss in logarithmic units is given by:

$$\mathrm{PL} = \begin{cases} -L - 35\log_{10}(d), & \text{if } d > d_1 \\ -L - 10\log_{10}\left(d_1^{1.5} d^2\right), & \text{if } d_0 < d \leq d_1 \\ -L - 10\log_{10}\left(d_1^{1.5} d_0^2\right), & \text{if } d < d_0 \end{cases}, \tag{3.3}$$

where d_0 and d_1 are the breakpoint distances of the three slope path loss model,

$$\begin{aligned} L = \; & 46.3 + 33.9\log_{10} f - 13.82\log_{10} h_{\mathrm{T}} - \\ & \left[1.11\log_{10} f - 0.7\right] h_{\mathrm{R}} + 1.56\log_{10} f - 0.8, \end{aligned} \tag{3.4}$$

f is the carrier frequency in MHz and h_{T} and h_{R} denote the transmitter and receiver antenna heights, respectively. Given the fact that the small-scale fading contribution to the entries of matrix $\mathbf{H}_{\mu}$ are i.i.d random variates, the channel matrix has full-rank with probability 1, and its rank is equal to the minimum value between N_T and N_R.

3.2.2 mm-Wave channel model

The measuring and modeling of mm-Wave channels have received considerable attention, leading to a solid understanding of how these channels differ from μ-Wave channels [37]. At mm-Wave frequencies, propagation mechanisms are different from those at μ-Waves. Indeed, path loss is much larger, while diffraction effects are practically negligible, thus implying that the typical range in cellular environments is usually not larger than 100 m, and the non-LOS component is mainly based on reflections. Moreover, signal blockages, due to the presence of macroscopic obstacles between the transmitter and the receiver, are much more frequent than those at μ-Wave frequencies. In order to catch these peculiarities, general consensus has been reached on the so-called clustered channel model [32], [38]–[41]. This model is based on the assumption that the propagation environment is made of N_{cl} scattering clusters, each of which contributes with N_{ray} propagation paths, plus a possibly present LOS component. Apart from the LOS component, the transmitter and the receiver are linked through single reflections on the N_{cl} scattering clusters.

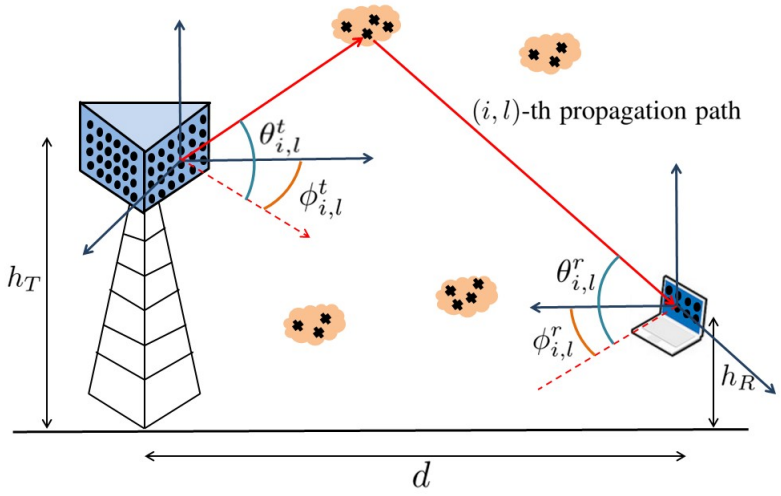

FIGURE 3.2: The considered reference scenario at mm-Wave frequencies.

The considered reference scenario for the clustered channel model at mm-Wave frequencies is reported in Fig. 3.2. Assuming again frequency-flat fading and focusing on a bi-dimensional model for the sake of simplicity, the baseband equivalent of the propagation channel is now represented by an $(N_R \times N_T)-$dimensional matrix expressed as:

$$\mathbf{H} = \gamma \sum_{i=1}^{N_{\text{cl}}} \sum_{\ell=1}^{N_{\text{ray}}} \alpha_{i,\ell} \sqrt{L(r_{i,\ell})} \mathbf{a}_r(\phi^r_{i,\ell}) \mathbf{a}_t^H(\phi^t_{i,\ell}) + \mathbf{H}_{\text{LOS}} \, . \tag{3.5}$$

In the above equation, neglecting for the moment $\mathbf{H}_{\text{LOS}}$, to be specified later, $\phi^r_{i,l}$ and $\phi^t_{i,l}$ are the azimuth angles of arrival and departure of the $\ell-$th ray in the $i-$th scattering cluster, respectively. The quantities $\alpha_{i,\ell}$ and $L(r_{i,\ell})$ are the complex path gain and the path loss associated to the (i,ℓ)-th propagation path (whose length is denoted by $r_{i,\ell}$), respectively. Following [33] the attenuation $L(r_{i,\ell})$ of the $(i,\ell)-$th path is written in logarithmic units as

$$L(r_{i,\ell}) = -20 \log_{10}\left(\frac{4\pi}{\lambda}\right) - 10n \left[1 - b + \frac{bc}{\lambda f_0}\right] \log_{10} r_{i,\ell} - X_\sigma \, , \tag{3.6}$$

with n the path loss exponent, X_σ the zero-mean, σ^2-variance Gaussian-distributed shadow fading term in logarithmic units, b a system parameter, c is the speed of light, and f_0 a fixed reference frequency, the centroid of all the frequencies represented by the path loss model. The values for all these parameters for each use-case scenario discussed in [33] (Urban Microcellular (UMi) Open-Square, UMi Street-Canyon, Indoor Hotspot (InH) Office, and InH Shopping Mall) are reported in Table 3.1.

TABLE 3.1: Parameters for Path Loss Model [33]

Scenario	Model Parameters
UMi Street Canyon LOS	$n = 1.98$, $\sigma = 3.1$ dB , $b = 0$
UMi Street Canyon NLOS	$n = 3.19$, $\sigma = 8.2$ dB , $b = 0$
UMi Open Square LOS	$n = 1.85$, $\sigma = 4.2$ dB , $b = 0$
UMi Open Square NLOS	$n = 2.89$, $\sigma = 7.1$ dB , $b = 0$
InH Indoor Office LOS	$n = 1.73$, $\sigma = 3.02$ dB , $b = 0$
InH Indoor Office NLOS	$n = 3.19$, $\sigma = 8.29$ dB $b = 0.06$, $f_0 = 24.2$ GHz
InH Shopping Mall LOS	$n = 1.73$, $\sigma = 2.01$ dB , $b = 0$
InH Shopping Mall NLOS	$n = 2.59$, $\sigma = 7.40$ dB $b = 0.01$, $f_0 = 39.5$ GHz

The complex gain $\alpha_{(i,\ell)} \sim \mathcal{CN}\left(0, \sigma^2_{\alpha_i}\right)$, with $\sigma^2_{\alpha_i} = 1$ [38]. Factors $\mathbf{a}_r(\phi^r_{i,\ell})$ and $\mathbf{a}_t(\phi^t_{i,\ell})$ represent the normalized receive and transmit array response vectors evaluated at the corresponding angles of arrival and departure; for an uniform linear array (ULA) with half-wavelength inter-element spacing

$$\mathbf{a}_t(\phi^t_{i,\ell}) = \frac{1}{\sqrt{N_T}}\left[1, e^{-j2\pi \sin\left(\phi^t_{i,\ell}\right)}, \ldots, e^{-j2\pi(N_T-1)\sin\left(\phi^t_{i,\ell}\right)}\right]^T , \tag{3.7}$$

and a similar expression can be also given for $\mathbf{a}_r(\phi^r_{i,\ell})$. Finally,

$$\gamma = \sqrt{\frac{N_R N_T}{\sum_{i=1}^{N_{\text{cl}}} N_{\text{ray},i}}} \tag{3.8}$$

is a normalization factor ensuring that the received signal power scales linearly with the product $N_R N_T$. Regarding the LOS component $\mathbf{H}_{\text{LOS}}$ in Eq. (3.5) the arrival and departure angles corresponding to the LOS link are denoted by ϕ^r_{LOS}, ϕ^t_{LOS} the LOS component is written as

$$\mathbf{H}_{\text{LOS}} = I_{\text{LOS}}(d)\sqrt{N_R N_T}e^{j\theta}\sqrt{L(d)}\mathbf{a}_r(\phi^r_{\text{LOS}}, \theta^r_{\text{LOS}})\mathbf{a}^H_t(\phi^t_{\text{LOS}}, \theta^t_{\text{LOS}}). \tag{3.9}$$

In the above equation, $\eta \sim \mathcal{U}(0, 2\pi)$, while $I_{\text{LOS}}(d)$ is a random variable that is 1 if a LOS link exists between transmitter and receiver and 0 otherwise. For the LOS probability the models in [33] is here assumed, that distinguishes between UMi and InH case. For the UMi scenarios the d_1/d_2 model is considered, with $d_1 = 20$ and $d_2 = 39$:

$$p^{\text{UMi}}(d) = \min\left(\frac{d_1}{d}, 1\right)\left(1 - e^{-\frac{d}{d_2}}\right) + e^{-\frac{d}{d_2}} , \tag{3.10}$$

and or the InH scenarios it is assumed:

$$p^{\mathrm{InH}}(d) = \begin{cases} 1 & d \leq 1.2, \\ e^{-\left(\frac{d-1.2}{4.7}\right)} & 1.2 < d \leq 6.5, \\ 0.32e^{-\left(\frac{d-6.5}{32.6}\right)} & d \geq 6.5. \end{cases} \tag{3.11}$$

A detailed description of all the parameters needed for the generation of sample realizations for the channel model in Eq. (3.5) is reported in [32]. Comparing the channel model in Eq. (3.5) for mm-Wave frequencies with the one in Eq. (3.1) for μ-Wave frequencies, it is immediately evident that the channel in Eq. (3.5) is a parametric channel model whose rank is tied to the number of clusters and reflectors contributing to the transmitter-receiver link. The next section will provide an accurate description of the implications that these two radically different channel models have on the architecture and on the attainable performance of massive MIMO multiuser wireless systems operating at μ-Wave and at mm-Wave frequencies.

3.3 Massive MIMO and doubly massive MIMO

The idea of a large scale antenna array was originally launched by Marzetta in his pioneering paper [4] with reference to BSs. The paper showed that in the limit of a large number of BS antennas small-scale fading effects vanish by virtue of channel hardening, and that channel vectors from the BS to the users tend to become orthogonal; consequently, plain CM beamforming at the BS permits serving several users on the same time-frequency resource slot with (ideally) no interference, and the only left impairment is imperfect channel estimates due to the fact that orthogonal pilots are limited and they must be re-used throughout the network (this is the so-called pilot contamination effect, discussed in detail in the Chapter 2 of the thesis). Reference [4] considered a system where MSs were equipped with just one antenna. Further studies have extended the massive MIMO idea at μ-Wave frequencies to the case in which the MSs have multiple antennas, but this number is obviously limited to few units. Indeed, at μ-Wave frequencies the wavelength is in the order of several centimeters, and it is thus difficult to pack many antennas on small-sized user devices. At μ-Waves, thus, massive MIMO just refers to BSs. Things are instead different at mm-Waves, wherein multiple antennas are necessary first and foremost to compensate for the increased path loss with respect to conventional sub-6 GHz frequencies.
Recalling the Friis transmission equation, the smaller wavelength λ directly increases the path-loss proportionally to λ^{-2}. This is due to the use of fixed-gain antennas whose effective area is proportional to λ^2. Hence, it can be overcome by using fixed-area antennas, which become increasingly directional with a gain proportional to λ^{-2}. The feasibility of communicating at a high rate in LOS, benefiting from the wide available bandwidth, also over long distances has been

exploited using high-gain directional antennas. Instead of deploying a huge array at one side of the link, the same signal-to noise ratio (SNR) can be achieved by deploying substantially smaller arrays at both sides. The beamforming gains are multiplied together, so instead of having 1000 BS antennas to serve single-antenna MSs, we can have 100 BS antennas and 10 MS antennas. This also opens the door to explore systems with massive arrays at both sides. This consideration leads to the concept of *doubly massive MIMO system* [29], [42], which is defined as a wireless communication system where the number of antennas grows large at both the transmitter and the receiver.
The douby massive MIMO system will be described in detail in the Chapter 4 of the thesis. Of course, there are a number of serious practical constraints – e.g., large power consumption, low efficiency of power amplifiers, hardware complexity, ADC and beamformer implementation – that currently prevent the feasibility of a user terminal equipped with hundreds of antennas. Mobile devices with a massive number of antennas thus will probably not be available in a few years, but, given the intense pace of technological progress, sooner or later they may become reality. As far as long-term forward-looking theoretical research is concerned, it is a concrete hypothesis that doubly-massive MIMO systems at mm-Waves will be a popular research topic for years to come.

3.4 Analog (beam-steering) beamforming optimality

One problem with massive MIMO systems is the cost and the complexity of needed hardware to efficiently exploit a so large number of antennas. If fully digital (FD) beamforming is to be made, as many radio frequency (RF) chains will be needed as the number of antennas; consequently, energy consumption will also grow linearly with the number of antennas. An evident concern is the implementation complexity of the digital baseband and analog/RF hardware. Technology scaling has fueled an impressive progress in wireless communication systems and it is essential to process many antenna signals. The flexibility offered by FD beamforming leads to the highest theoretically achievable performance, while hybrid analog-digital beamforming schemes are explored to enable hardware reuse over antenna paths. However, neither the digital processing nor the data converters are a complexity hurdle, although those are the stages where hybrid beamforming primarily induces simplifications. The high-speed interconnect that is the bottleneck in the realization of integrated systems. Recently, lower complexity architectures have been proposed, encompassing, for instance, 1-bit quantization of the antenna outputs [43] and hybrid analog/digital beamforming structures [38], [41], [44] wherein an RF beamforming matrix (whose entries operate as simple phase shifters) is cascaded to a reduced-size digital beamformer. In [45] it has been shown that if the number of RF chains

is twice the multiplexing order, the hybrid beamformer is capable of implementing any FD beamformer. Now, while at μ-Waves the use of hybrid beamformer brings an unavoidable performance degradation, at mm-Waves something different happens in the limiting regime of large number of antennas by virtue of the different propagation mechanisms. Indeed, the channel matrix in Eq. (3.5) can be compactly re-written as:

$$\mathbf{H} = \gamma \sum_{i=1}^{N} \alpha_i \mathbf{a}_r(\phi_i^r)\mathbf{a}_t^H(\phi_i^t)\,, \tag{3.12}$$

lumping the coefficients α_i into the path-loss term, and grouping the two summations over the clusters and the rays in just one summation, with N being the number of propagation paths from the transmitter to the receiver. Given the continuous random location of the scatterers, the set of arrival angles will be different with probability 1, i.e., there is a zero probability that two distinct scatterers will contribute to the channel with the same departure and arrival angles. Since, for a large number of antennas, $\mathbf{a}_x^H(\phi_p^x)\mathbf{a}_x(\phi_p^x) \rightarrow 0$, provided that $\phi_p^x \neq \phi_q^x$, $x = \{r,t\}$, it is possible to conclude that for large N_T , vectors $\mathbf{a}_t(\phi_i^t)$ for all $i = 1,\ldots,N$ converge to an orthogonal set, and, similarly, for large N_R, the vectors $\mathbf{a}_r(\phi_i^r)$ for all $i = 1,\ldots,N$ converge to an orthogonal set as well. Accordingly, in the doubly massive MIMO regime, the array response vectors $\mathbf{a}_t(\cdot)$ and $\mathbf{a}_r(\cdot)$ become the left and right singular vectors of the channel matrix, i.e., the channel representation in Eq. (3.12) coincides with the singular-value-decomposition of the channel matrix. Under this situation, purely analog (beam-steering) beamforming becomes optimal. Otherwise stated, two main consequences are pointed out. First, in a single-user link, the channel eigendirections associated to the largest eigenvalues are just the beam-steering vectors corresponding to the arrival and departure angles and associated with the predominant scatterers. This suggests that pre-coding and post-coding beamforming simply require pointing a beam towards the predominant scatterer at the transmitter and at the receiver respectively. Second, in a multiuser environment, assuming that the links between the several users and the BS involve separate scatterers and different sets of arrival and departure angles[3], beam-steering analog beamforming automatically results in no-cochannel interference (in the limiting regime of infinite number of antennas) since the beams pointed towards different users tend to become orthogonal.

[3]This is a quite reasonable assumption for sufficiently spaced mobile user locations

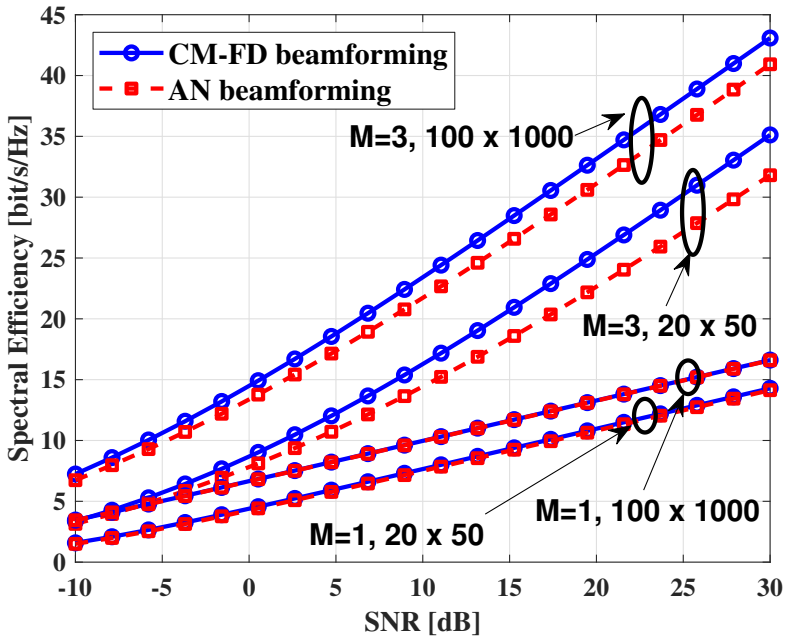

FIGURE 3.3: Spectral efficiency of a mm-Wave MIMO wireless link vs. received SNR for CM-FD beamforming and AN (beam-steering) beamforming, for two different values of the number of transmit and receive antennas and of the multiplexing order of the system.

Fig. 3.3 provides some experimental evidence of the above statements. It is considered a single-user MIMO link at mm-Waves; the carrier frequency is 73 GHz, the transmitting antenna height is 15 m, while the receiving antenna height is 1.65 m. All the parameters needed for the generation of the mm-Wave channel matrix in Eq. (3.5) are the ones reported in [32] for the "open square model". Fig. 3.3 shows the system spectral efficiency measured in bit/s/Hz, versus the received SNR, and it compares the performance of the CM FD beamforming and the analog (AN) beam-steering beamforming. With CM beamforming the pre-coding and post-coding beamformers are the left and singular eigenvectors of the channel matrix in Eq. (3.5) associated to the M largest eigenvalues respectively; with AN beamforming, instead, the pre-coding and post-coding beamformers are simply the array responses corresponding to the departure and arrival angles associated to the M dominant scatterers respectively. From the figure it is seen that AN beamforming achieves practically the same performance as CM beamforming for multiplexing order $M = 1$, even in the case of not-so-large number of antennas, while there is a small gap for $M = 3$; this gap is supposed to get reduced as the number of antennas increases.

3.5 Rank of the channel not increasing with N_T and N_R

At μ-Wave frequencies, the i.i.d. assumption for the small-scale fading component of the channel matrix $\mathbf{H}_\mu$ guarantees that with probability 1 the matrix has rank equal to $\min(N_T, N_R)$. Consequently, as long as the rich-scattering environment assumption holds and the number of degrees of freedom of the radiated and scattered fields is sufficiently high [46], the matrix rank increases linearly with the number of antennas. At mm-Wave frequencies, instead, the validity of the channel model in Eq. (3.5) directly implies that, including the LOS component, the channel has at most the rank $N_{\text{cl}}N_{\text{ray}} + 1$, since it is expressed as the sum of $N_{\text{cl}}N_{\text{ray}} + 1$ rank-1 matrices. This rank is clearly independent of the number of transmit and receive antennas, so, mathematically, as long as $\min(N_T, N_R) > N_{\text{cl}}N_{\text{ray}} + 1$, increasing the number of antennas has no effect on the channel rank. However, it is also suggested that, for an increasing number of antennas, the directive beams become narrower and more scatterers can be resolved, which implies that the channel rank increases (even though probably not linearly) with the number of antennas. However, this is still a conjecture that would need experimental validation. With respect to the number of antennas, the described different behavior of the channel rank has a profound impact on the multiplexing capabilities of the channel. Indeed, for μ-Wave systems, the increase in the channel rank leads to an increase of the multiplexing capabilities of the channel; on the other hand, the multiplexing capabilities depend on the number of scatterers in the propagation environment in mm-Wave systems, while the number of antennas just contributes to the increase of the received power that can increase proportionally to the product $N_T N_R$.

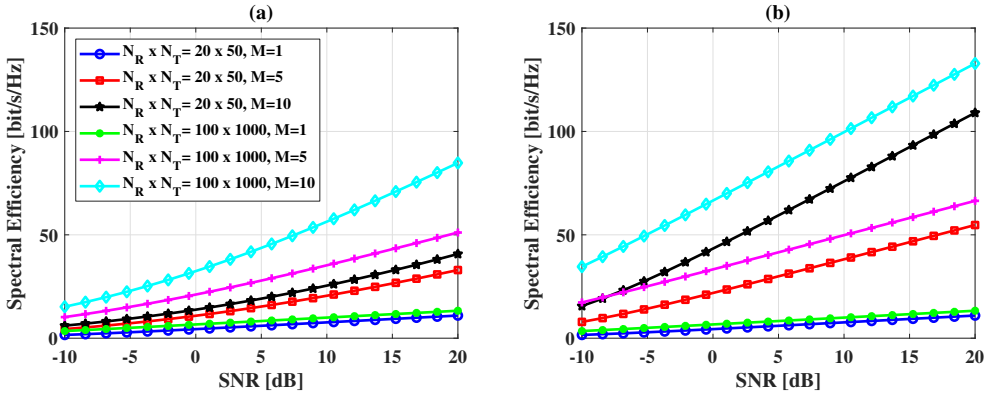

FIGURE 3.4: In (a) spectral efficiency vs. received SNR for an mm-Wave channel varying the number of transmit and receive antennas and multiplexing order, and in (b) spectral efficiency vs. received SNR for a μ-Wave channel varying the number of transmit and receive antennas and multiplexing order.

Fig. 3.4 provides experimental evidence of such a different behavior. The figure shows the system spectral efficiency for mm-Wave and μ-Wave wireless MIMO

links, for two different values of the number of receive and transmit antennas, and for three different values of the multiplexing order M. The parameters of the mm-Wave channel are the same as those in Fig. 3.3. Regarding the μ-Wave channel, a carrier frequency equal to 1.9 GHz is considered and the standard deviation of the shadow fading σ_{sh} is taken equal to 8 dB, while the parameters of the three-slope path loss model in Eq. (3.3) are $d_0 = 50$ m and $d_1 = 100$ m. It is clearly seen from Fig. 3.4 that the μ-Wave channel has larger multiplexing capabilities than the mm-Wave channel; the gap between the two scenarios is mostly emphasized for the large values of M and for $N_R \times N_T$=100 x 1000.

3.6 Channel estimation

The number of channel coefficients grows linearly with the number of antennas at the BS and MS. To have an approximate idea of the computational burden, consider a system with 200 BS antennas and 20 spatially multiplexed single-antenna MSs. Consider OFDM with 1024 subcarriers and channels that are constant over 12 subcarriers. There are $3.4\,10^5$ complex scalar coefficients, which amounts to $6.8\,10^6$ estimates each second if a channel coherence time of 50 ms is assumed. These numbers increase if there are more antennas, more subcarriers, and/or shorter coherence time.

3.6.1 μ-Wave channel estimation complexity

At μ-Wave, there is generally multi-path propagation caused by a multitude of scattering clusters. The channel coefficients are correlated across antennas, but this can only be utilized to marginally improve the estimation quality, at the cost of substantially higher complexity. Nevertheless, the estimation can be conveniently implemented in hardware and the estimation overhead is small when operating in TDD mode and exploiting channel reciprocity to only send uplink pilots. Anyway, in μ-Wave massive MIMO systems, channel estimation is a rather difficult and resource-consuming task, since it requires the separate estimation of each entry of the matrix $\mathbf{H}_\mu$. It thus follows that in a multiuser system with K users equipped with N_R antennas each, the number of parameters to be estimated is KN_RN_T, where N_T denotes the number of antennas at the BS. The attendant computational complexity needed to perform channel estimation is a growing function of the number of used antennas. The increase of the number of antennas N_R at the mobile devices has a direct impact on the network capacity. Indeed, let τ_c denote the duration (in discrete samples) of the channel coherence time and τ_p the length (again in discrete samples) of the pilot sequences used on the uplink for channel estimation; since the length of pilot sequences must be a fraction (typically no more than 1/2) of the channel coherence length, and since the use of orthogonal pilots across users requires that $KN_R \leq \tau_p < \tau_c$, it is readily seen that there is a physical bound on the maximum number of users

and the number of transceiver antennas at the mobile device. Such a bound is the main underlying motivation for the fact that a considerable share of the available literature on massive MIMO systems at μ-Waves focuses on the case of single-antenna mobile devices, and with $N_R = 1$, the number of users K can be taken larger. In order to increase the number of supported users, pseudo-orthogonal pilots with low cross-correlation are used, even though this leads to the well-known pilot contamination problem that, as discussed in Chapter 2, is the ultimate performance limit in μ-Wave massive MIMO systems [4].

3.6.2 mm-Wave channel estimation complexity

At mm-Wave, the channel can potentially be parameterized (considering a phase-synchronized array with a known angular array response as reported in Eq. (3.5)) because it consists of a (potential) LOS path and few one-bounce reflections. Instead of estimating the individual entries of the channel matrix $\mathbf{H}$, a few angular channel coefficients can be estimated to acquire the entire channel, leading to greatly reduced complexity. When a single data-stream is to be sent, it suffices to estimate the dominant angle-of-arrival/departure, but also reflections can be taken into account. However, if hybrid beamforming is used, the phase-shifters create a very directional "vision" and only channel components in that direction can be estimated. To discover new MSs, track channel variations, or keep the connection when the LOS path is blocked, beam-sweeping is needed (i.e., the channel must be estimated in many different directions to identify the preferable ones). This procedure increases the overhead from CSI acquisition, which grows with the number of antennas. While TDD operation is preferable at μ-Wave massive MIMO, in mm-Wave bands frequency-division-duplex (FDD) may be equally good since the channel-describing angular parameters are reciprocal over a wide bandwidth. Based on this consideration, the computational complexity of the channel estimation schemes at mm-Waves may be smaller than that at μ-Waves.

3.6.3 Performance comparison

Channel estimation for mm-Wave frequencies is a research track that is currently under development, whereas for μ-Waves this is a rather mature area. Among the several existing approaches to perform channel estimation at mm-Waves, the most considered ones rely either on compressed sensing or on subspace methods. As an example, reference [47] shows that at mm-Waves, for increasing number of antennas, the most significant components of the received signal lie in a low-dimensional subspace due to the limited angular spread of the reflecting clusters. This low-dimensionality feature can be exploited in order to obtain channel estimation algorithms based on the sampling of only a small subset rather than of the whole number of antenna elements. Consequently, channel

estimation can be performed using a reduced number (with respect to the number of receive antennas) of required RF chains and A/D converters at receiver front-end. Reference [48], instead, develops subspace-based channel estimation methods exploiting channel reciprocity in time-division-duplex (TDD) systems, using the well-known Arnoldi iteration and explicitly taking into account the adoption of hybrid analog/digital beamforming structures at the transmitter and at the receiver. Subspace methods are particularly attractive in those situations where it is of interest to estimate the principal left and right singular eigenvectors of the channel matrix **H**, which, in the doubly massive MIMO regime, are well-approximated by the array response vectors corresponding to the dominant scatterers. As done in [49], applying fast subspace estimation algorithms such as the Oja's one [50], the dominant channel eigenvectors can be directly obtained by the sample estimate of the data covariance matrix, with no need to directly estimate the whole channel matrix **H**.

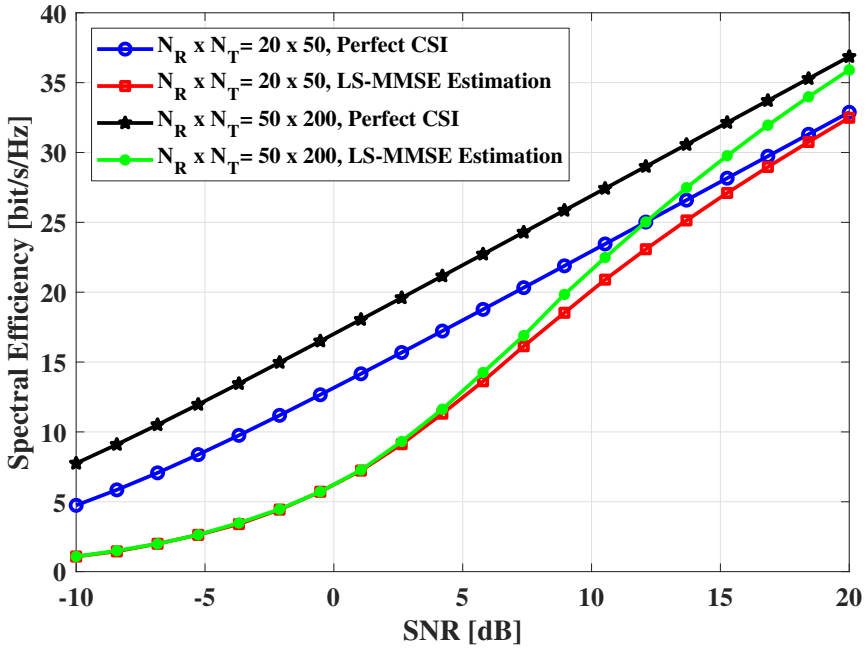

FIGURE 3.5: Spectral efficiency vs. received SNR with perfect CSI and imperfect CSI, with LS-MMSE algorithm for the estimation of μ-Wave channel. The multiplexing order is 3.

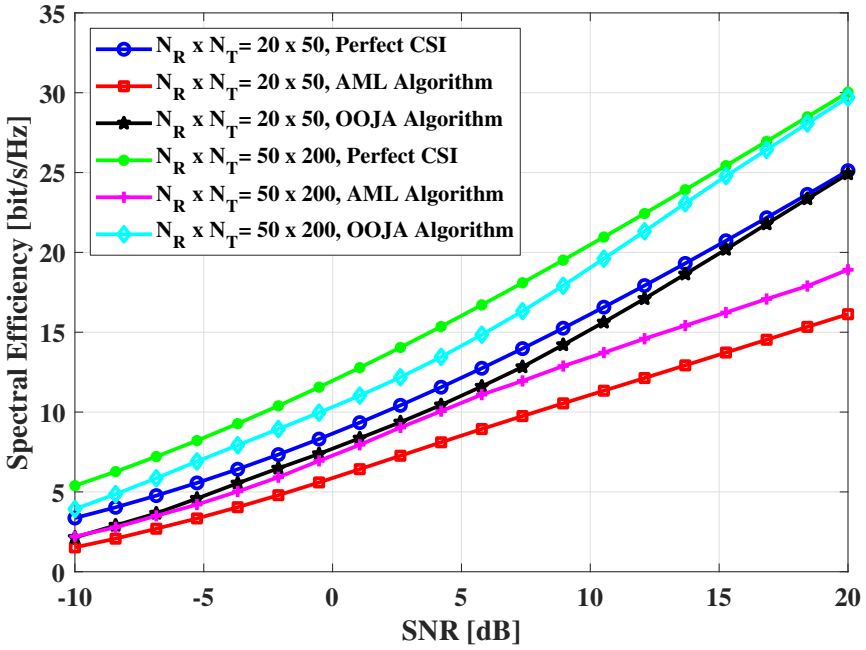

FIGURE 3.6: Spectral efficiency vs. received SNR with perfect CSI and imperfect CSI, with AML algorithm and OOJA algorithm for the estimation of mm-Wave channel. The multiplexing order is 3.

Figs. 3.5 and 3.6 show numerical results concerning channel estimation at μ-Wave and at mm-Wave channel frequencies. In particular, both figures report the spectral efficiency vs. the received SNR for two different antenna configurations and by contrasting the case of perfect channel state information (CSI) with the case in which the channel is estimated based on training pilots. In both figures a single-user MIMO link is considered, and channel estimation is carried out assuming that each transmit antenna sends an orthogonal pilot. The number of signaling intervals devoted to channel estimation coincides with the number of transmit antennas. Note that this is the minimum possible duration in order to be able to send orthogonal pilots. Channel estimation at μ-Wave frequencies (Fig. 3.5) is made using the linear minimum mean square errors (LMMSE) criterion ([51]), while at mm-Wave frequencies (Fig. 3.6) the approximate maximum likelihood (AML) algorithm of [47] and the orthogonal Oja (OOJA) algorithm [49] are used. Comparing the figures, it is clearly seen that the gap between the case of estimated channel and the case of perfect CSI is smaller at mm-Wave frequencies, especially when the OOJA algorithm is considered. Conversely, this gap is larger at μ-Waves, and it grows with the dimension of the user antenna arrays. This behavior can be intuitively explained by virtue of the parametric form of the mm-Wave channel model in Eq. (3.5), which permits the development of efficient channel estimation algorithms.

3.7 Pilot contamination effect

Pilot contamination is the ultimate disturbance in massive MIMO systems operating at μ-Waves. As already discussed above, the impossibility to have a number of orthogonal pilots larger than the number of signaling intervals devoted to channel estimation leads to the use of pseudo-orthogonal, low cross-correlation sequences. Accordingly, in a massive MIMO system, when the MSs transmit their own pilot sequences in the uplink training phase to enable channel estimation at the BSs, every BS learns the channel from the intended MS, and also small pieces of the channels from the other MSs using pilots that are correlated to the one used by the intended MS. This phenomenon, in turn, causes a saturation in the achieved SINR both in the downlink and in the uplink. The deceitful nature of pilot contamination in massive MIMO systems was discussed by Marzetta in his landmark paper [4] and since then, many authors have deeply investigated its effects and proposed strategies to counterbalance its effects [10], [52], [53]. All of these papers deal with the case of a μ-Wave massive MIMO system. Pilot contamination at mm-Wave frequencies is instead a much less-studied topic (some initial results are reported in [54]). This is in part due to the fact that massive MIMO at mm-Waves is a more recent research topic than massive MIMO at μ-Waves. On the other hand, it may be envisioned that pilot contamination may be less critical at mm-Waves than it has revealed at μ-Waves, mainly for the short-range nature of mm-Wave links. In particular, while the range of μ-Wave links can be in the order of thousands of meters, the range for mm-Wave links will be more than one order of magnitude smaller, due to the increased path loss and a larger relevance of signal blockages. Mm-Wave frequencies will be used for short-range communications in small cells, which, by nature, usually serve a smaller number of users than conventional micro-cells and macro-cells. Therefore, on one hand, the signals transmitted by the MSs during uplink training fade rapidly with the distance, and thus they should not be a serious impairment to surrounding BSs learning the channel from their intended MSs; on the other hand, the reduced number of users in each cell will lead to a less severe shortage of orthogonal pilots. The results in [54] seem to confirm such increased resilience of mm-Waves to the pilot contamination problem.

3.8 Antenna diversity/selection procedures

The i.i.d. nature of the fast fading component in the MIMO channel matrix at μ-Waves in Eq. (3.1) leads to a monotonic increase of the diversity order with the number of antennas. In particular, an $N_R \times N_T$ channel brings a diversity order equal to $N_R \times N_T$, thus implying that the average error probability decreases to a zero, in the limit of large SNR, as $\text{SNR}^{-N_R N_T}$. Such a diversity order can be attained through a simple antenna selection procedure by picking the transmit and receive antennas corresponding to the entry with the largest magnitude in

the channel matrix $\mathbf{H}_\mu$. Looking at this fact from a different perspective, it is possible to recall the well-known probability result stating that the maximum of a set of positive i.i.d. random variables taking value in the interval $[0, +\infty)$, becomes unbounded with probability 1 as the cardinality of the set diverges. As a consequence, for increasing number of antennas, the probability of observing a very large entry in the channel matrix rapidly increases. The open literature is rich of studies exploiting this peculiarity of μ-Wave MIMO channels and proposing diversity techniques based on antenna selection procedures (e.g., [55] and [56]). At mm-Waves, instead, given the parametric channel model in Eq. (3.5), a different behavior is observed. In particular, the entries of the matrix channel have no longer an i.i.d. component, and this implies that the maximum of the magnitudes of the entries of $\mathbf{H}$ grows at a much reduced pace. As a consequence, diversity techniques using antenna selection procedures are less effective. As an experimental evidence of this fact, the parameter η is defined as the ratio between the largest squared magnitude among the entries of the channel, and the average squared magnitude

$$\eta = N_R N_T \frac{\max_{i,j} \left| \left(\widehat{\mathbf{H}} \right)_{(i,j)} \right|^2}{\mathrm{tr}\left(\widehat{\mathbf{H}} \widehat{\mathbf{H}}^H \right)}, \tag{3.13}$$

where $\widehat{\mathbf{H}} = \mathbf{H}$ in the case of mm-Wave channel and $\widehat{\mathbf{H}} = \mathbf{H}_\mu$ in the case of μ-Wave channel.

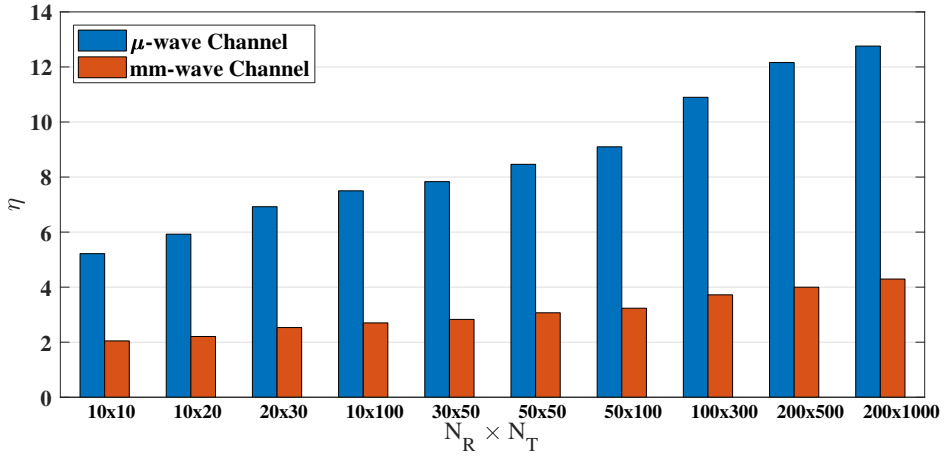

FIGURE 3.7: Values of the performance measure η for several antenna array sizes for the mm-Wave and μ-Wave channels.

Fig. 3.7 reports η for different values of $N_R \times N_T$, and for both the μ-Wave and mm-Wave channel models. The larger η is, the more unbalanced are the magnitudes of the entries of the channel matrix, since η basically measures how far is the largest entry in $\mathbf{H}$ from the average magnitude. Fig. 3.7, shows that the

parameter η is in general an increasing function of the number of antenna elements, but it grows much more rapidly in the case of μ-Wave channels.

Chapter 4

Doubly massive MIMO mm-Wave systems

Referred articles published in the *IEEE Transactions on Green Communications and Networking* 2018 and in the*Proc. of 2016 IEEE Global Communications Conference (GLOBECOM)*.

4.1 Introduction

The use of frequency bands at mm-Waves for cellular communications is among the most striking technological innovations brought by 5G and beyond-5G wireless networks[4] [1]. For conventional sub-6 GHz cellular systems it has been shown that equipping a BS with a very large (> 100) number of antennas, the massive MIMO technique [4], [25], [57], significantly increases the network capacity, mainly due to the capability of serving several users on the same frequency slot with nearly orthogonal vector channels. In the massive MIMO literature, while the number of antennas at the BS grows large, the user device is usually assumed to have only one or very few antennas: indeed, at sub-6 GHz frequencies the wavelength is in the order of several centimeters, thus making it difficult to pack many antennas on small-sized user devices. When moving to mm-Wave, however, the wavelength gets reduced, and, at least in principle, a large number of antennas can be mounted not only on the BS, but also on the user device. As an example, at a carrier frequency of 30 GHz the wavelength is 1 cm, and for a planar antenna array with $\lambda/2$ spacing, more than 180 antennas can be placed in an area as large as a standard credit card. This leads to the concept of *doubly massive MIMO system* [29], [58], that is defined as a wireless communication system where the number of antennas grows large at both the transmitter and the receiver.
While there are certainly a number of serious practical constraints – e.g., large power consumption, low efficiency of power amplifiers, hardware complexity, ADC and beamformer implementation – that currently prevent the feasibility of

[4]Strictly speaking, mm-Waves is a term that historically refers to the range of frequencies above 30 GHz, in the recent literature about future wireless networks the term is used to refer to frequency above-6 GHz, in contrapposition to the usual cellular frequencies located below 6 GHz.

an user terminal equipped with a very large number of antennas, it is on the other hand believed that these are just technological issues that will be solved or worked around in the near future, and thus this chapter of the thesis presents results on the doubly massive MIMO regime for wireless systems operating at mm-Wave. When considering MIMO architectures, and in particular massive MIMO ones, hardware complexity and energy consumption issues make the use of conventional FD beamforming, which requires one RF chain for each antenna element, rather prohibitive; as a consequence, recent research efforts have been devoted towards devising suboptimal, lower complexity, beamforming structures [59]. In particular, hybrid (HY) beamforming structures have been proposed, with a limited number (much smaller than the number of antenna elements) of RF chains. The paper [60] analyzes the achievable rate for a MU-MIMO system with HY pre-coding and limited feedback; it is therein shown that, for the case of single-path (i.e., rank-1) channels, HY pre-coding structures achieve a spectral efficiency very close to that of a FD beamformer. In [45], it is shown that a HY beamformer with twice as many RF chains as transmitted data streams may exactly mimic a FD beamformer; the analysis, which neglects energy efficiency issues, is however limited to either a single-user MIMO system or a MU-MIMO system with single-antenna receivers. The paper [61] proposes a new low-complexity post-coding structure, based on switches rather than on analog phase shifters; the performance of this new structure is evaluated in a rather simple scenario, i.e., single-user MIMO system with a limited number of transmit and receive antennas. The paper [62] focuses on sub-6 GHz frequencies and introduces a novel post-coding structure made of fixed (rather than tunable) phase shifters and of switches, under the assumption that the receiver is equipped with a large array, while the transmitters have only one antenna. In [44], the authors considers five different low-complexity decoding structures, all based on the use of phase shifters and switches, and provide an analysis of the achievable spectral efficiency along with estimates of the energy consumption of the proposed structures. The paper, however, does not analyze the system energy efficiency (i.e., the ratio of the achievable rate to the consumed power [63]), and focuses only on the receiver omitting a similar analysis for the transmitter implementation. The paper [64] considers the issue of energy efficiency in a downlink massive MIMO mm-Wave systems by deriving an energy-efficient HY beamformer; however, the paper considers the case in which the user terminals are equipped with just one antenna, and this is a key assumption that is exploited to solve the considered optimization problems. In [65], a consumed power model for components designed for 60 GHz is given, and a comparison between FD beamforming, 1-bit ADC, and analog beamforming is given.

This chapter of the thesis focuses on both the ASE and the GEE of several pre-coding and post-coding structures, ranging from the FD beamformers, to their

HY analog-digital implementations, to the lower complexity purely AN beamforming structures. It also proposes extensions to the considered scenario of recently proposed beamforming structures such as the one based on switches [61] and the one based on fixed phase shifters and switches [62]. Putting emphasis on the large number of antennas regime, asymptotic formulas of the ASE and of the GEE are derived with reference to some of these post-coding structures. Special emphasis, at the analysis stage, is also given to the purely AN (beam-steering) beamformer, that can be considered as a promising solution given its extremely low complexity. While the results on the ASE confirm, as expected, that FD beamforming achieves better performance than lower-complexity structures, things are a little bit more involved and surprising when considering the GEE. Indeed, here, the relative ranking of the several low-complexity structures strongly depends on the adopted power consumption model for amplifiers, phase shifters, switches, etc., and our results show that, using recent power models, FD beamforming may be the most energy-efficient solution. The chapter of the thesis also studies the system GEE as a function of the transmitted power, and numerically finds the optimal value for the transmitted power (around the value 0 dBW in the downlink); going beyond this point increasing the transmit power is not convenient from an energy-efficient point of view since it yields a limited increase in the network throughput at the price of a strong increase of the consumed energy.
The remainder of this chapter is organized as follows. In Section 4.2 the system model is described. Section 4.3 contains the description of the several considered beamforming structures, while in Section 4.4 asymptotic formulas for the system ASE and GEE of two FD beamformers, in the limit of large number of antennas are derived, both for the uplink and downlink. Section 4.5 is entirely devoted to the exposition of asymptotic results for the purely analog beam-steering beamformers, while in Section 4.6 extensive numerical results are discussed.

4.2 System model

This chapter of the thesis focuses on a single-cell MU-MIMO system wherein one BS communicates, on the same frequency slot, with several mobile users. As in Chapter 3, the parameter N_T denotes the number of transmit antennas at the BS, and N_R denotes the number of receive antennas at the user's device[5].

[5]For the sake of simplicity all the mobile receivers are assumed to have the same number of antennas; however, this hypothesis can be easily relaxed.

4.2.1 Clustered channel model

The popular narrowband clustered mm-Wave channel model is assumed to hold [38]–[40], [66]. As detailed in Section 3.2.2, the baseband equivalent of the propagation channel between the transmitter and the generic receiver is thus represented by Eq. (3.5).

4.2.2 Transmitter and receiver processing

- **Downlink**: Assume that M denotes the number of data symbols sent to each user in each signalling interval[6], and $\mathbf{x}_k$ is the M-dimensional vector of the data symbols intended for the k-th user; the discrete-time signal transmitted by the BS can be expressed as the N_T-dimensional vector $\mathbf{s}_T = \sum_{k=1}^{K} \mathbf{Q}_k \mathbf{x}_k$, with $\mathbf{Q}_k$ the $(N_T \times M)$-dimensional pre-coding matrix for the k-th user. The signal received by the generic k-th user is expressed as the following N_R-dimensional vector

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{s}_T + \mathbf{w}_k \,, \tag{4.1}$$

  with $\mathbf{H}_k$ representing the clustered channel (modeled as in Eq. (3.5)) from the BS to the k-th user and $\mathbf{w}_k$ is the N_R-dimensional additive white Gaussian noise with zero-mean i.i.d. entries with variance σ_n^2. Denoting by $\mathbf{D}_k$ the $(N_R \times M)$-dimensional post-coding matrix at the k-th user device, the following M-dimensional vector is finally obtained:

$$\mathbf{r}_k = \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_k \mathbf{x}_k + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_\ell \mathbf{x}_\ell + \mathbf{D}_k^H \mathbf{w}_k \,. \tag{4.2}$$

- **Uplink**: In an uplink scenario, K denotes the number of users simultaneously transmitting to the BS in the same frequency band, and M again denotes the number of data symbols sent by each user in each signalling interval. Letting now $\mathbf{x}_k$ be the M-dimensional vector of the data symbols from the k-th user, the discrete-time signal transmitted by the k-th user device is expressed as the N_T-dimensional vector $\mathbf{s}_k = \mathbf{Q}_k \mathbf{x}_k$, with $\mathbf{Q}_k$ the $(N_T \times M)$-dimensional pre-coding matrix for the k-th user. The signal received by the BS is expressed as the following N_R-dimensional vector

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{Q}_k \mathbf{x}_k + \mathbf{w} \,, \tag{4.3}$$

  with $\mathbf{H}_k$ representing now the channel from the k-th user to the BS and $\mathbf{w}$ the N_R-dimensional additive white Gaussian noise with zero-mean i.i.d. entries with variance σ_n^2. Assuming, for the sake of simplicity, single-user

[6] Otherwise stated, the BS transmits in each time-frequency slot MK data symbols.

processing at the BS, a soft estimate of the symbols from the k-th user is obtained as

$$\widehat{\mathbf{x}}_k = \mathbf{D}_k^H \mathbf{y} = \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_k \mathbf{x}_k + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{D}_k^H \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{x}_\ell + \mathbf{D}_k^H \mathbf{w} \,, \tag{4.4}$$

with $\mathbf{D}_k$ the $(N_R \times M)$-dimensional post-coding matrix for the k-th user symbols. Now, depending on the choice of the pre-coding and post-coding matrices $\mathbf{Q}_k$ and $\mathbf{D}_k$, several transceiver structures can be conceived. These will be illustrated later in the next section.

4.2.3 Performance measures

Two performance measures will be considered: the ASE and the GEE. The ASE is measured in [bit/s/Hz], while the GEE is measured in [bit/Joule] [63]. Assuming Gaussian data symbols in (4.2), the ASE for the downlink case is [67]

$$\text{ASE} = \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T}{KM} \mathbf{R}_{D,\bar{k}}^{-1} \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_k \mathbf{Q}_k^H \mathbf{H}_k^H \mathbf{D}_k \right| \,, \tag{4.5}$$

wherein $\mathbf{I}_M$ is the identity matrix of order M, P_T is the BS transmit power, and, according again to the signal model (4.2) $\mathbf{R}_{D,\bar{k}}$ is the covariance matrix of the overall disturbance seen on the downlink by the k-th user receiver, i.e.,

$$\mathbf{R}_{D,\bar{k}} = \sigma_n^2 \mathbf{D}_k^H \mathbf{D}_k + \frac{P_T}{MK} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_\ell \mathbf{Q}_\ell^H \mathbf{H}_k^H \mathbf{D}_k \,. \tag{4.6}$$

For the uplink, instead, the k-th user ASE is [67]

$$\text{ASE}_k = \log_2 \left| \mathbf{I}_M + \frac{P_{T,k}}{M} \mathbf{R}_{U,\bar{k}}^{-1} \mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_k \mathbf{Q}_k^H \mathbf{H}_k^H \mathbf{D}_k \right| \,, \tag{4.7}$$

$\forall k = 1, \ldots, K$, wherein $P_{T,k}$ is the k-th user transmit power, and the overall disturbance covariance matrix, according to the signal model in (4.4), is now written as[7]

$$\mathbf{R}_{U,\bar{k}} = \sigma_n^2 \mathbf{D}_k^H \mathbf{D}_k + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \frac{P_{T,\ell}}{M} \mathbf{D}_k^H \mathbf{H}_\ell \mathbf{Q}_\ell \mathbf{Q}_\ell^H \mathbf{H}_\ell^H \mathbf{D}_k \,. \tag{4.8}$$

Regarding the GEE, on the downlink it is defined as

$$\text{GEE} = \frac{W\text{ASE}}{\eta P_T + P_{\text{TX,c}} + K P_{\text{RX,c}}} \,, \tag{4.9}$$

where W is the system bandwidth, $P_{\text{TX,c}}$ is the amount of power consumed by the BS circuitry, $P_{\text{RX,c}}$ is the amount of power consumed by the mobile user's

[7] Note that the power budget, both at the BS and at the user's transmitters, is assumed to be uniformly divided among the data streams, although power allocation could be easily performed.

device circuitry, and $\eta > 1$ is a scalar coefficient modelling the power amplifier inefficiency. Note that, differently from what happens in the most part of existing studies on energy efficiency for cellular communications (see, for instance, references of [63]), the GEE definition (4.9) includes here the power consumed both at the BS and at the mobile user's devices.
For the uplink scenario, instead the GEE of the k-th user is

$$\mathrm{GEE}_k = \frac{W\mathrm{ASE}_\mathrm{k}}{\eta P_{T,k} + P_{\mathrm{TX,c}}}\,, \tag{4.10}$$

where $P_{\mathrm{TX,c}}$ is now the amount of power consumed by the k-th mobile device circuitry. Notice that in this case the power consumed by the BS is not included in the definition; this is a customary choice when defining energy efficiency in the uplink in order to have a performance measure that can be used for non-cooperative resource allocation procedures [63], [68].

4.3 Beamforming structures

In the following, some beamforming pre-coding and post-coding structures are detailed, along with details on their power consumption. The section mainly focuses on the downlink, although the uplink case can be treated with minor modifications.

4.3.1 Channel-matched, fully-digital (CM-FD) beamforming

Let $\mathbf{H}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{V}_k^H$ denote the singular-value-decomposition (SVD) of the matrix $\mathbf{H}_k$, and assume, without loss of generality, that the diagonal entries of $\mathbf{\Lambda}_k$ are sorted in descending order. The column vectors $\mathbf{u}_{k,i}$ and $\mathbf{v}_{k,i}$ denote the i-th column of the matrices $\mathbf{U}_k$ and $\mathbf{V}_k$, respectively. The k-th user pre-coding and post-coding matrices $\mathbf{Q}_k^{\mathrm{CM-FD}}$ and $\mathbf{D}_k^{\mathrm{CM-FD}}$ are chosen as the columns of the matrices $\mathbf{V}_k$ and $\mathbf{U}_k$, respectively, corresponding to the M largest entries in the eigenvalue matrix $\mathbf{\Lambda}_k$, i.e.,

$$\begin{array}{rcl} \mathbf{Q}_k^{\mathrm{CM-FD}} & = & \left[\mathbf{v}_{k,1}\ \mathbf{v}_{k,2}\ \ldots\ \mathbf{v}_{k,M}\right], \\ \mathbf{D}_k^{\mathrm{CM-FD}} & = & \left[\mathbf{u}_{k,1}\ \mathbf{u}_{k,2}\ \ldots\ \mathbf{u}_{k,M}\right], \end{array} \tag{4.11}$$

$\forall k = 1, \ldots, K$. The CM-FD beamforming is optimal in the interference-free case, and tends to be optimal in the case in which the number of antennas at the transmitter grows large. The considered FD pre-coding architecture requires a baseband digital precoder that adapts the M data streams to the N_T transmit antennas; then, for each antenna there is a digital-to-analog-converter (DAC), an RF chain and a power amplifier (PA). At the receiver, a low noise amplifier (LNA), an RF chain, an analog-to-digital converter (ADC) is required for each antenna, plus a baseband digital combiner that combines the N_R outputs of ADC

to obtain the soft estimate of the M trasmitted symbols. The amount of power consumed by the transmitter circuitry can thus be expressed as

$$P_{\text{TX,c}} = N_T \left(P_{\text{RFC}} + P_{\text{DAC}} + P_{\text{PA}}\right) + P_{\text{BB}} \,, \tag{4.12}$$

and the amount of power consumed by the receiver circuitry can be expressed as

$$P_{\text{RX,c}} = N_R \left(P_{\text{RFC}} + P_{\text{ADC}} + P_{\text{LNA}}\right) + P_{\text{BB}} \,. \tag{4.13}$$

In the above equations, $P_{\text{RFC}} = 40$ mW [44] is the power consumed by the single RF chain, $P_{\text{DAC}} = 110$ mW [69] is the power consumed by each DAC, $P_{\text{ADC}} = 200$ mW [44] is the power consumed by each ADC, $P_{\text{PA}} = 16$ mW [70] is the power consumed by each PA, $P_{\text{LNA}} = 30$ mW [44] is the power consumed by each LNA, and P_{BB} is the amount of power consumed by each baseband beamformers; assuming a CMOS implementation we have a power consumption of 243 mW [71]. The values of the power consumed by the each ADC present high variability in the current literature [44]. A conservative value is chosen since the literature does not refer to commercial products and these values might be too optimistic with respect to the final products. The values of the power consumption of each device considered in this section are summarized in Table 4.1.

TABLE 4.1: Power consumption of each device

Name	**Value**	**Device**	**Reference**
P_{RFC}	40 mW	RF chain	[44]
P_{DAC}	110 mW	DAC	[69]
P_{ADC}	200 mW	ADC	[44]
P_{PA}	16 mW	PA	[44]
P_{LNA}	30 mW	LNA	[44]
P_{BB}	243 mW	Baseband beamformer	[71]
P_{PS}	19.5 mW	Phase shifter	[72]
P_{element}	27 mW	Element of the pased array	[70]
P_{SW}	5 mW	Switch	[44]
$P_{\text{PS}}^{\text{fixed}}$	1 mW	Constant phase shifter	-

4.3.2 Partial zero-forcing, fully digital (PZF-FD) beamforming

Zero-forcing pre-coding nulls interference at the receiver through the constraint that the k-th user pre-coding be such that the product $\mathbf{H}_\ell \mathbf{Q}_k = \mathbf{0}_{N_T \times M}$ for all $\ell \neq k$. In order to avoid a too severe noise enhancement, a partial zero-forcing approach is adopted here, namely the columns of the pre-coding matrix $\mathbf{Q}_k$ are required to be orthogonal to the M (the number of transmitted data-streams to each user) right eigenvectors of the channel $\mathbf{H}_\ell$ corresponding to the largest eigenvalues of $\mathbf{H}_\ell$, for all $\ell \neq k$. In this way, the precoder orthogonalizes

only to a $M(K-1)$-dimensional subspace and nulls the most significant part of the interference. Formally, the precoder $\mathbf{Q}_k^{\mathrm{PZF-FD}}$ is obtained as the projection of the CM-FD precoder $\mathbf{Q}_k^{\mathrm{CM-FD}}$ onto the orthogonal complement of the subspace spanned by the M dominant right eigenvectors of the channel matrices $\mathbf{H}_1, \ldots, \mathbf{H}_{k-1}, \mathbf{H}_{k+1}, \ldots, \mathbf{H}_K$. The post-coding matrix is instead obtained as $\mathbf{D}_k^{\mathrm{PZF-FD}} = (\mathbf{H}_k \mathbf{Q}_k^{\mathrm{PZF-FD}})^+$. Since the PZF-FD beamforming requires a FD post-coding, its power consumption is the same as that of the CM-FD beamformer.

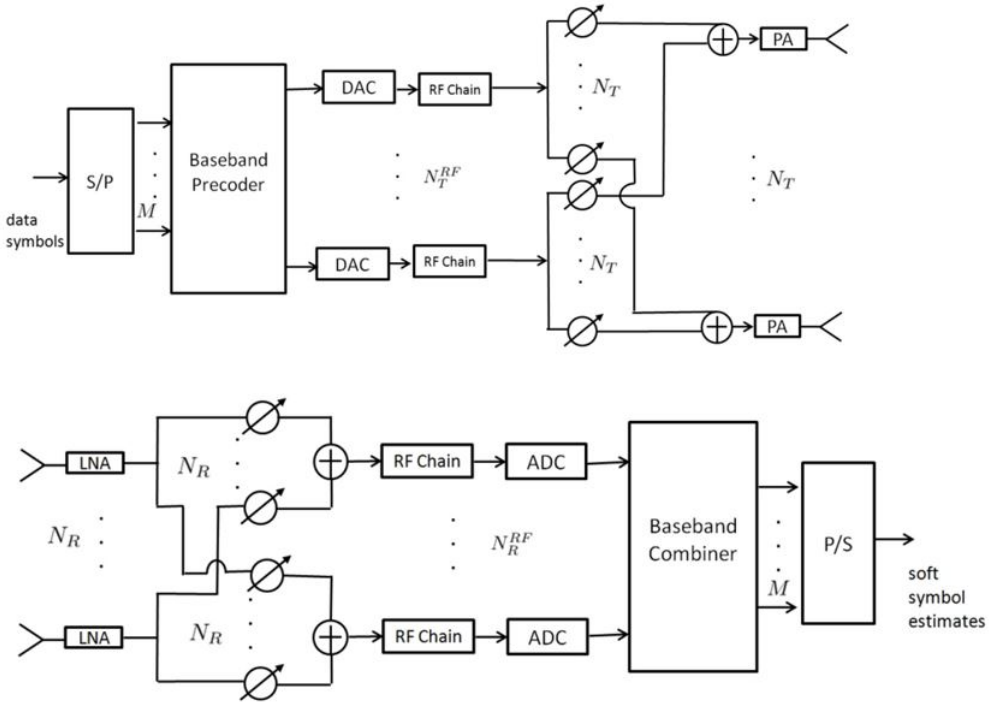

FIGURE 4.1: Block-scheme of a transceiver with HY digital/analog beamforming.

4.3.3 Channel-matched, hybrid (CM-HY) beamforming

In order to avoid the same number of RF chains as the number of antennas, HY beamforming architectures have been proposed; in particular, denoting by N_T^{RF} and N_R^{RF} the number of RF chains available at the transmitter and at the receiver, respectively, the k-th user pre-coding and post-coding matrices are decomposed as follows:

$$\mathbf{Q}_k^{\mathrm{CM-HY}} = \mathbf{Q}_k^{\mathrm{RF}} \mathbf{Q}_k^{\mathrm{BB}}, \quad \mathbf{D}_k^{\mathrm{CM-HY}} = \mathbf{D}_k^{\mathrm{RF}} \mathbf{D}_k^{\mathrm{BB}}. \tag{4.14}$$

In the above decomposition, the matrices $\mathbf{Q}_k^{\mathrm{RF}}$ and $\mathbf{D}_k^{\mathrm{RF}}$ have dimension $(N_T \times N_T^{\mathrm{RF}})$ and $(N_R \times N_R^{\mathrm{RF}})$, respectively, and their entries are constrained to have constant (unit) norm (i.e., they are implemented through a network of phase-shifters[8]); the matrices $\mathbf{Q}_k^{\mathrm{BB}}$ and $\mathbf{D}_k^{\mathrm{BB}}$, instead, have dimension $(N_T^{\mathrm{RF}} \times M)$ and

[8]The case of quantized phase-shifts is also considered in the literature, but it is neglected here for the sake of simplicity.

Algorithm 1 Block Coordinate Descent for Subspace Decomposition Algorithm for Hybrid Beamforming

1: Initialize $I_{\max}$ and set $i = 0$
2: Set arbitrary $\mathbf{Q}_{\mathrm{RF},0}$ and $\mathbf{D}_{\mathrm{RF},0}$
3: **repeat**
4: Update $\mathbf{Q}_{\mathrm{BB,i+1}} = \left(\mathbf{Q}_{\mathrm{RF,i}}^H \mathbf{Q}_{\mathrm{RF,i}}\right)^{-1} \mathbf{Q}_{\mathrm{RF,i}}^H \mathbf{Q}^{\mathrm{opt}}$
and $\mathbf{D}_{\mathrm{BB,i+1}} = \left(\mathbf{D}_{\mathrm{RF,i}}^H \mathbf{D}_{\mathrm{RF,i}}\right)^{-1} \mathbf{D}_{\mathrm{RF,i}}^H \mathbf{D}^{\mathrm{opt}}$
5: Set $\phi_i = \mathbf{Q}^{\mathrm{opt}} \mathbf{Q}_{\mathrm{BB,i+1}}^H \left(\mathbf{Q}_{\mathrm{BB,i+1}} \mathbf{Q}_{\mathrm{BB,i+1}}^H\right)^{-1}$
and $\psi_i = \mathbf{D}^{\mathrm{opt}} \mathbf{D}_{\mathrm{BB,i+1}}^H \left(\mathbf{D}_{\mathrm{BB,i+1}} \mathbf{D}_{\mathrm{BB,i+1}}^H\right)^{-1}$
6: Update $\mathbf{Q}_{\mathrm{RF,i}} = \frac{1}{\sqrt{N_T}} e^{j\phi_i}$
and $\mathbf{D}_{\mathrm{RF,i}} = \frac{1}{\sqrt{N_R}} e^{j\psi_i}$
7: Set $i = i + 1$
8: **until** convergence or $i = I_{\max}$

$(N_R^{\mathrm{RF}} \times M)$, respectively, and their entries are unconstrained complex numbers. A block-scheme of the architecture of the HY transceiver is depicted in Fig. 4.1. Now, designing an HY beamformer is tantamount to finding expressions for the matrices $\mathbf{Q}_k^{\mathrm{RF}}, \mathbf{Q}_k^{\mathrm{BB}}, \mathbf{D}_k^{\mathrm{RF}}$, and $\mathbf{D}_k^{\mathrm{BB}}$, so that the desired beamformers reported in Eq. (4.14) are approximated. For the CM-HY beamforming, the desired beamformers are the PZF-FD matrices, and their approximation is realized by using the block coordinate descent for subspace decomposition algorithm [48], [73], briefly reported in Algorithm 1, with $\mathbf{Q}^{\mathrm{opt}} = \mathbf{Q}^{\mathrm{CM-FD}}$ and $\mathbf{D}^{\mathrm{opt}} = \mathbf{D}^{\mathrm{CM-FD}}$.
The number of RF chains at the BS is KM, while at the mobile terminal it is M. The amount of power consumed by the transmitter circuitry is [44]:

$$P_{\mathrm{TX,c}} = N_T^{\mathrm{RF}} \left(P_{\mathrm{RFC}} + P_{\mathrm{DAC}} + N_T P_{\mathrm{PS}}\right) + N_T P_{\mathrm{PA}} + P_{\mathrm{BB}} \,, \tag{4.15}$$

and the amount of power consumed by the receiver circuitry is:

$$P_{\mathrm{RX,c}} = N_R^{\mathrm{RF}} \left(P_{\mathrm{RFC}} + P_{\mathrm{ADC}} + N_R P_{\mathrm{PS}}\right) + N_T P_{\mathrm{LNA}} + P_{\mathrm{BB}} \,. \tag{4.16}$$

Numerical values for the above quantities have already been given, except that for P_{PS}, the power consumed by each phase shifters, that is assumed to be 19.5 mW as in [72].

4.3.4 Partial zero-forcing, hybrid (PZF-HY) beamforming

Similarly to what has been described in the previous subsection, also the PZF beamformers may be approximated through HY architectures. In this case, expressions for the matrices $\mathbf{Q}_k^{\mathrm{RF}}, \mathbf{Q}_k^{\mathrm{BB}}, \mathbf{D}_k^{\mathrm{RF}}$, and $\mathbf{D}_k^{\mathrm{BB}}$ are to be found, so that the PZF-FD beamforming matrices are approximated as closely as possible. Also in this case the block coordinate descent for subspace decomposition algorithm

can be used and the hybrid beamformers are evaluated following Algorithm 1, with $\mathbf{Q}^{\text{opt}} = \mathbf{Q}^{\text{PZF-FD}}$ and $\mathbf{D}^{\text{opt}} = \mathbf{D}^{\text{PZF-FD}}$. Again the number of RF chains at the BS is KM, while at the mobile terminal is M. The amount of power consumed by the transmitter circuitry of th PZF-HY beamformers is the same as that consumed by the CM-HY ones.

4.3.5 Fully analog (AN) beam-steering beamforming

Fully analog beamforming requires that the entries of the pre-coding and post-coding matrices have a constant norm. Here, it is considered an even simpler structure by introducing a further constraint and assuming that the columns of matrices $\mathbf{Q}_k$ and $\mathbf{D}_k$ are unit-norm beam-steering vectors, i.e., the generic column of an N-dimensional beamformer is

$$\mathbf{a}(\phi) = \frac{1}{\sqrt{N}}[1 \; e^{-jkd\sin\phi} \; \ldots \; e^{-jkd(N-1)\sin\phi}] \; . \tag{4.17}$$

Focusing on the generic k-th user, the columns of the matrix $\mathbf{Q}_k^{\text{AN}}$ are chosen as the array responses corresponding to the departure angles in the channel model (3.5) associated to the M dominant paths. A similar choice is made for $\mathbf{D}_k^{\text{AN}}$, whose columns contain the array responses corresponding to the M arrival angles associated to the M dominant paths. In order to avoid self-interference, a further constraint is added in the choice of the dominant paths to ensure that the angles of departure (arrival) of the selected paths are spaced of at least 5 deg. Note that for large values of N_T and N_R the array responses of the transmitter and receiver, corresponding to the departure and arrival angles associated to the dominant propagation path, become concident with dominant right and left singular vectors of the channel. This implies that the AN beamforming structure (4.17) tends to become optimal. The amount of power consumed by the transmitter circuitry is:

$$P_{\text{TX,c}} = N_T^{\text{RF}} \left(P_{\text{RFC}} + N_T P_{\text{element}} + P_{\text{DAC}}\right) \; , \tag{4.18}$$

and the amount of power consumed by the receiver circuitry is:

$$P_{\text{RX,c}} = N_R^{\text{RF}} \left(P_{\text{RFC}} + N_R P_{\text{element}} + P_{\text{ADC}}\right) \; , \tag{4.19}$$

where $P_{\text{element}} = 27$ mW [70] is the power consumed by each element of the phased array.

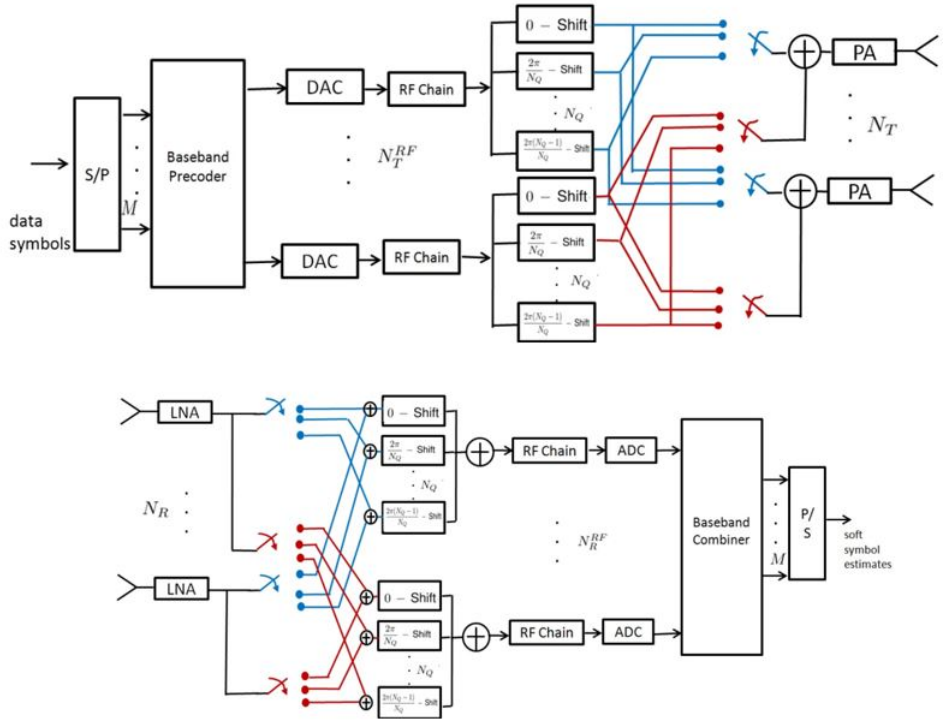

FIGURE 4.2: Block-scheme of a transceiver where beamforming is implemented with switches and N_Q constant phase shifters per RF chain.

4.3.6 Beamforming based on switches and fixed phase shifters (SW+PHSH)

The considered structure, depicted in Fig. 4.2, builds upon [62], wherein a massive MIMO combiner is proposed based on the use of switches and fixed (i.e., not tunable) phase shifters. The scheme in Fig. 4.2 extends the structure of [62] by including also the pre-coding design. The design of the SW+PHSH beamformers is detailed in the following. The (i,ℓ)-entry of the pre-coding matrix is in the form $\left(\mathbf{Q}^{\text{SW+PHSH}}\right)_{(i,\ell)} = e^{j\phi_{i,\ell}}$, where the phase $\phi_{i,\ell}$ can take only discrete quantized values. It is thus a unitary module entry with a quantized phase that is obtained substituting the phase of corresponding entry of the pre-coding matrix intended to be synthesized with the nearest quantizated phase, taken from the set $\left\{\frac{2(q-1)\pi}{N_Q}, q = 1,\ldots,N_Q\right\}$. The procedure for the design of the SW+PHSH pre-coder is detailed in Algorithm 2, with $\mathbf{Q}^{\text{opt}} = \mathbf{Q}^{\text{PZF-FD}}$. A similar reasoning is followed for the entries of the post-coding matrix $\mathbf{D}^{\text{SW+PHSH}}$.

The number of quantized phases will be $N_Q = 8$, and that the number of RF chains is KM at the BS and M at the users' devices. The number of fixed phase shifters per RF chain is N_Q, along with N_T^{RF} and N_R^{RF} switches per antenna at the transmitter and at the receiver, respectively. The amount of power consumed by the transmitter circuitry is:

$$P_{\text{TX,c}} = N_T^{\text{RF}}\left(P_{\text{RFC}} + P_{\text{DAC}} + N_Q P_{\text{PS}}^{\text{fixed}}\right) + N_T\left(N_T^{\text{RF}} P_{\text{SW}} + P_{\text{PA}}\right) + P_{\text{BB}}\,, \quad (4.20)$$

Algorithm 2 Design of the SW+PHSH pre-coder

1: Initialize $\vartheta_q = \frac{2(q-1)\pi}{N_Q}, q = 1, \ldots, N_Q$
2: **for** $i = 1, \ldots, N_T$ **do**
3: **for** $\ell = 1, \ldots, M$ **do**
4: $\phi_{i,\ell} = \underset{q=1,\ldots,N_Q}{\operatorname{argmin}} \angle \mathbf{Q}^{\text{opt}} - \vartheta_q$
5: Set $\left(\mathbf{Q}^{\text{SW+PHSH}}\right)_{(i,\ell)} = e^{j\phi_{i,\ell}}$
6: **end for**
7: **end for**

and the amount of power consumed by the receiver circuitry is:

$$P_{\text{RX,c}} = N_R^{\text{RF}} \left(P_{\text{RFC}} + P_{\text{ADC}} + N_Q P_{\text{PS}}^{\text{fixed}}\right) + N_R \left(N_R^{\text{RF}} P_{\text{SW}} + P_{\text{LNA}}\right) + P_{\text{BB}} \,. \tag{4.21}$$

In the above equations, $P_{\text{SW}} = 5$ mW [44] is the power consumed by the single switch, and $P_{\text{PS}}^{\text{fixed}}$ is the power consumed by the constant phase shifter; this term is of course lower than the power consumed by a tunable phase shifter, and is set to 1 mW.

4.3.7 Switch-based (SW) beamforming

A beamforming structure exclusively based on the use of switches is reported in [61]. Once again, N_T^{RF} and N_R^{RF} denote the number of RF chains at the transmitter and at the receiver, respectively, and it is assumed that there are N_T^{RF} switches at the transmitter and N_R^{RF} at the receiver that select the antennas using the Minimum Frobenius Norm (MFN) algorithm in [61]. The design of the switch-based beamformers is detailed in the following. The pre-coding matrix is in the form $\mathbf{Q}^{\text{SW}} = \mathbf{S}\mathbf{Q}^{\text{BB}}$ where $\mathbf{S}$ is a $\left(N_T \times N_T^{\text{RF}}\right)$-dimensional matrix with columns that have exactly one position containing the value 1 and the other entries in the matrix are 0, and $\mathbf{Q}^{\text{BB}}$ is the $N_T^{\text{RF}} \times M$-dimensional baseband pre-coding matrix. It can be thus shown that the matrix $\mathbf{Q}^{\text{SW}}$ contains non-zero N_T^{RF} rows corresponding to the N_T^{RF} rows of the pre-coding matrix intended to be synthesized with the largest norm. The procedure for the design of the SM pre-coder is detailed in Algorithm 3. A similar reasoning is followed for the entries of the post-coding matrix $\mathbf{D}^{\text{SW}}$.

Again, the number of RF chains in the BS is assumed to be equal to KM, while at the mobile terminal it is equal to M. The amount of power consumed by the transmitter circuitry is:

$$P_{\text{TX,c}} = N_T^{\text{RF}} \left(P_{\text{RFC}} + P_{\text{DAC}} + P_{\text{SW}}\right) + N_T^{\text{RF}} P_{\text{PA}} + P_{\text{BB}} \,, \tag{4.22}$$

and the amount of power consumed by the receiver circuitry is:

$$P_{\text{RX,c}} = N_R^{\text{RF}} \left(P_{\text{RFC}} + P_{\text{ADC}} + P_{\text{SW}}\right) + N_R^{\text{RF}} P_{\text{LNA}} + P_{\text{BB}} \,. \tag{4.23}$$

Algorithm 3 Design of the SW pre-coder

1: Initialize $\mathbf{S} = \mathbf{0}_{N_T \times N_T^{\mathrm{RF}}}$
2: **for** $\ell = 1, \ldots, N_T^{\mathrm{RF}}$ **do**
3: Set $\mathbf{q}_\ell = \left(\mathbf{Q}^{\mathrm{BB}}\right)_{(:,\ell)}$
4: Compute the vector $|\mathbf{q}_\ell|^2$, containing the square module of the entries of $\mathbf{q}_\ell$
5: Denote as i^* the position of the maximum entry of $|\mathbf{q}_\ell|^2$
6: Set $(\mathbf{S})_{(i^*,\ell)} = 1$
7: **end for**
8: $\mathbf{Q}^{\mathrm{SW}} = \mathbf{S}\mathbf{Q}^{\mathrm{BB}}$

4.4 Asymptotic ASE and GEE analysis for the CM-FD and PZF-FD beamformers for large number of antennas

4.4.1 CM-FD beamforming, downlink

Focusing on the downlink, in the large number of antennas regime, making the assumption that the set of arrival and departure angles across clusters and users are different with probability 1, it readily follows from the SVD expression of the channel that $\mathbf{D}_k^H \mathbf{H}_k \mathbf{Q}_\ell \to \mathbf{\Lambda}_{k,M} \mathbf{V}_{k,M} \mathbf{Q}_\ell$, whenever $k \neq \ell$, where $\mathbf{\Lambda}_{k,M}$ is an $(M \times M)-$dimensional diagonal matrix containing the M largest eigenvalues (denoted by $\lambda_{k,1}, \ldots, \lambda_{k,M}$) of the channel matrix $\mathbf{H}_k$ and $\mathbf{V}_{k,M}$ is an $(N_T \times M)$-dimensional matrix containing the columns of $\mathbf{V}_k$ associated to the eigenvalues in $\mathbf{\Lambda}_{k,M}$. Using the above limiting values, the asymptotic ASE in Eq. (4.5) can be expressed as

$$\mathrm{ASE} \approx \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T}{KM} \left(\sigma_n^2 \mathbf{I}_M + \frac{P_T}{MK} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{\Lambda}_{k,M} \mathbf{V}_{k,M}^H \mathbf{Q}_\ell \mathbf{Q}_\ell^H \mathbf{V}_{k,M} \mathbf{\Lambda}_{k,M}^H \right)^{-1} \mathbf{\Lambda}_{k,M} \mathbf{\Lambda}_{k,M}^H \right| . \tag{4.24}$$

In order to explicitly show the dependence of the above formula on the number of antennas, note that the squared moduli of the eigenvalues $\lambda_{k,i}$ depend linearly on the product $N_T N_R$. Otherwise stated, the following holds: $\lambda_{k,q} = \sqrt{N_T N_R} \widetilde{\lambda}_{k,q}$, $\forall k, q$, with $\widetilde{\lambda}_{k,q}$ normalized eigenvalues independent of the number of transmit and receive antennas. Using this last relation, and denoting by $\{\mu_{k,q}\}_{q=1}^{M}$ the eigenvalues of the matrix $\sum_{\ell=1,\ell\neq k}^{K} \mathbf{\Lambda}_{k,M} \mathbf{V}_{k,M}^H \mathbf{Q}_\ell \mathbf{Q}_\ell^H \mathbf{V}_{k,M} \mathbf{\Lambda}_{k,M}^H$, straightforward manipulations lead to the following alternative expression for the ASE is:

$$\mathrm{ASE} \approx \sum_{k=1}^{K} \sum_{q=1}^{M} \log_2 \left(1 + N_R N_T \frac{\frac{P_T}{MK} |\widetilde{\lambda}_{k,q}|^2}{\sigma_n^2 + \frac{P_T}{MK} \mu_{k,q}} \right) . \tag{4.25}$$

Eq. (4.25) confirms that with the clustered channel model scenario increasing the number of antennas does not provide additional degrees of freedom but just a SINR-gain proportional to the product $N_T N_R$. Now, Eqs. (4.9), (4.25) can be

used to obtain an expression for the asymptotic GEE, i.e.,

$$\text{GEE} \approx \frac{\sum_{k=1}^{K}\sum_{q=1}^{M} W \log_2\left(1 + N_R N_T \frac{\frac{P_T}{MK}|\widetilde{\lambda}_{k,q}|^2}{\sigma_n^2 + \frac{P_T}{MK}\mu_{k,q}}\right)}{\eta P_T + P_{\text{TX,c}} + K P_{\text{RX,c}}} . \tag{4.26}$$

An interesting problem is the GEE maximization with respect to the transmitted power, the number of transmit antennas and the number of receive antennas[9]. While global GEE maximization with respect to N_T, N_R and P_T may be cumbersome, it is worth noting that a sequential iterative algorithm, wherein at each iteration maximization with respect to one parameter only is performed, can be easily conceived. Indeed, it is easily seen that the fraction in (4.26) is the ratio of a concave function (with respect to the single variables N_T, N_R, and P_T) over a linear one, and, thus, Dinkelbach's algorithm may be readily applied to maximize the ratio [68]. Rigorousy speaking, the variables P_T is continuous, while N_T, N_R are integer variables, so that the numerator of (4.26) is not a concave function with respect to N_T and N_R. However, this problem is often addressed in literature considering N_T and N_R as continuous variables, solving the corresponding problem and finally concluding that the quasiconcavity of the problem implies that the optimal value for each variable is attained at one of the two closest integers [74], [75]. Further details on this for the sake of brevity are not provided. However, in the section on the numerical results plots of the GEE versus P_T, providing an insight on the range of transmit power values that maximize the system energy efficiency, will be reported.

4.4.2 CM-FD beamforming, uplink

Similar conclusions can be also drawn for the uplink scenario. Note that in this case N_T denotes the number of antennas on the user's device and N_R denotes the BS array size. For large number of antennas, now it holds $\mathbf{D}_k^H \mathbf{H}_\ell \mathbf{Q}_\ell \approx \mathbf{D}_k^H \mathbf{U}_{\ell,M} \mathbf{\Lambda}_{\ell,M}$ and the k-th user ASE is

$$\text{ASE}_k \approx \log_2 \left| \mathbf{I}_M + \frac{P_{T,k}}{M} \left(\sigma_n^2 \mathbf{I}_M + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \frac{P_{T,\ell}}{M} \mathbf{D}_k^H \mathbf{U}_{\ell,M} \mathbf{\Lambda}_{\ell,M} \mathbf{\Lambda}_{\ell,M}^H \mathbf{U}_{\ell,M}^H \mathbf{D}_k \right)^{-1} \mathbf{\Lambda}_{k,M} \mathbf{\Lambda}_{k,M}^H \right| . \tag{4.27}$$

Substituting (4.27) into (4.10) it is finally possible to obtain an asymptotic expression for the GEE of the generic k-th user.

4.4.3 PZF-FD beamforming, downlink

For PZF-FD beamforming, the product $\mathbf{H}_k \mathbf{Q}_\ell$ is an all-zero matrix whenever $k \neq \ell$. As a consequence, $\mathbf{R}_{D,\bar{k}} \approx \sigma_n^2 \mathbf{I}_M$, and the asymptotic ASE can be shown

[9] Recall that $P_{\text{TX,c}}$ and $P_{\text{RX,c}}$ depend linearly on N_T and N_R, respectively.

to be written as[10]

$$\text{ASE} \approx \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T}{KM\sigma_n^2} \mathbf{\Lambda}_{k,M} \mathbf{\Lambda}_{k,M}^H \right| = \sum_{k=1}^{K} \sum_{q=1}^{M} \log_2 \left(1 + \frac{P_T}{MK} \frac{|\lambda_{k,q}|^2}{\sigma_n^2} \right) . \tag{4.28}$$

Using the normalized eigenvalues $\widetilde{\lambda}_{k,q} = \lambda_{k,q} / \sqrt{N_T N_R}$, $\forall k, q$, the following equivalent expression is obtained:

$$\text{ASE} \approx \sum_{k=1}^{K} \sum_{q=1}^{M} \log_2 \left(1 + \frac{N_T N_R P_T}{MK} \frac{|\widetilde{\lambda}_{k,q}|^2}{\sigma_n^2} \right) . \tag{4.29}$$

The GEE is now written as:

$$\text{GEE} \approx \frac{\sum_{k=1}^{K} \sum_{q=1}^{M} \log_2 \left(1 + \frac{N_T N_R P_T}{MK} \frac{|\widetilde{\lambda}_{k,q}|^2}{\sigma_n^2} \right)}{\eta P_T + P_{\text{TX,c}} + K P_{\text{RX,c}}} , \tag{4.30}$$

and, also in this case, Dinkelbach's algorithm can be successfully applied to perform alternative maximization of the GEE with respect to N_T, N_R and the transmit power P_T.

4.4.4 PZF-FD beamforming, uplink

Exploiting the fact that $\mathbf{D}_k^H \mathbf{H}_\ell$ is zero whenever $k \neq \ell$, the asymptotic ASE for the k-th user becomes:

$$\text{ASE}_k \approx \sum_{q=1}^{M} \log_2 \left(1 + \frac{N_T N_R P_{T,k}}{M} \frac{|\widetilde{\lambda}_{k,q}|^2}{\sigma_n^2} \right) . \tag{4.31}$$

Substituting (4.31) into (4.10) it is finally possible to obtain an asymptotic expression for the GEE of the generic k-th user.

4.5 Asymptotic ASE and GEE analysis for the beam-steering AN beamformers for large number of antennas

4.5.1 AN beamforming, downlink

The case of AN pre-coding and post-coding is now considered. As a preliminary step to the analysis, it is convenient to recall that the ULA response in Eq. (4.17) is a unit-norm vector, and that the inner product between two ULA responses of length P and corresponding to incidence angles ϕ_1 and ϕ_2 is written as

$$f_P(\phi_1, \phi_2) \triangleq \mathbf{a}^H(\phi_1) \mathbf{a}(\phi_2) = \frac{1}{P} \frac{1 - e^{jkd(\sin\phi_1 - \sin\phi_2)P}}{1 - e^{jkd(\sin\phi_1 - \sin\phi_2)}} . \tag{4.32}$$

[10] This is an asymptotic expression since the noise enhancement effect is neglected (that is a decreasing function of N_T) induced by the nulling of the interference.

The above inner product, that is denoted by $f_P(\phi_1, \phi_2)$, has a magnitude that, for large P, vanishes as $1/P$, whenever $\phi_1 \neq \phi_2$. Let us now write the channel matrix for user k as

$$\mathbf{H}_k = \gamma_k \sum_{i=1}^{N} \alpha_{k,i} \mathbf{a}_r(\phi_{i,k}^r) \mathbf{a}_t^H(\phi_{i,k}^t) = \gamma_k \mathbf{A}_{k,r} \mathbf{L}_k \mathbf{A}_{k,t}^H \,, \tag{4.33}$$

namely the path-loss term has been lumped into the coefficients $\alpha_{\cdot,\cdot}$, and the summation over the clusters and the rays has been compressed in just one summation, with $N = N_{\text{cl}} N_{\text{ray}}$. Additionally, $\mathbf{A}_{k,r}$ is an $(N_R \times N)$-dimensional matrix containing on its columns the vectors $\mathbf{a}_r(\phi_{1,k}^r), \ldots, \mathbf{a}_r(\phi_{N,k}^r)$, $\mathbf{L}_k = \text{diag}(\alpha_{1,k}, \ldots, \alpha_{N,k})$, and $\mathbf{A}_{k,t}$ is an $(N_T \times N)$-dimensional matrix containing on its columns the vectors $\mathbf{a}_t(\phi_{1,k}^t), \ldots, \mathbf{a}_t(\phi_{N,k}^t)$[11]. It is also assumed, with no loss of generality, that the paths are sorted in decreasing magnitude order, i.e., $|\alpha_{1,k}| \geq |\alpha_{2,k}| \geq \ldots \geq |\alpha_{N,k}|$. In the following analysis, it is assumed that there are no collisions between arrival and departure angles across users, an assumption that is usually verified unless there are very close users. For the downlink scenario, the analog post-coding and pre-coding matrices are written as

$$\begin{aligned} \mathbf{D}_k &= [\mathbf{a}_r(\phi_{1,k}^r), \ldots, \mathbf{a}_r(\phi_{M,k}^r)] \,, \\ \mathbf{Q}_k &= [\mathbf{a}_t(\phi_{1,k}^t), \ldots, \mathbf{a}_t(\phi_{M,k}^t)] \,, \end{aligned} \tag{4.34}$$

$\forall k$, and they are actually submatrices of $\mathbf{A}_{k,r}$ and $\mathbf{A}_{k,t}$, respectively. Define now the following $(M \times N)$-dimensional matrices: $\mathbf{F}_{k,\ell,M}^r \triangleq \mathbf{D}_k^H \mathbf{A}_{\ell,r}$ and $\mathbf{F}_{k,\ell,M}^t \triangleq \mathbf{Q}_k^H \mathbf{A}_{\ell,t}$ Note that the (m,n)-th entry of the matrix $\mathbf{F}_{k,\ell,M}^t$ is $f_{N_T}(\phi_{m,k}^t, \phi_{n,\ell}^t)$, while the (m,n)-th entry of the matrix $\mathbf{F}_{k,\ell,M}^r$ is $f_{N_R}(\phi_{m,k}^r, \phi_{n,\ell}^r)$. Equipped with this notation, the ASE in (4.5) can be now expressed as follows:

$$\text{ASE} = \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T}{KM} \gamma_k^2 \mathbf{R}_{D,\bar{k}}^{-1} \mathbf{F}_{k,k,M}^r \mathbf{L}_k \mathbf{F}_{k,k,M}^{t\,H} \mathbf{F}_{k,k,M}^t \mathbf{L}_k^* \mathbf{F}_{k,k,M}^{r\,H} \right| \,, \tag{4.35}$$

with

$$\mathbf{R}_{D,\bar{k}} = \sigma_n^2 \mathbf{D}_k^H \mathbf{D}_k + \frac{P_T}{MK} \gamma_k^2 \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \mathbf{F}_{k,k,M}^r \mathbf{L}_k \mathbf{F}_{\ell,k,M}^{t\,H} \mathbf{F}_{\ell,k,M}^t \mathbf{L}_k^* \mathbf{F}_{k,k,M}^{r\,H} \,. \tag{4.36}$$

In order to have an asymptotic expression of Eq. (4.35) for a large number of antennas, it can be noted that the $(M \times N)$-dimensional matrix $\mathbf{F}_{k,\ell,M}^r$ is such that (a) for $k \neq \ell$ all its entries have a norm that for large N_R vanishes as $1/N_R$; while (b) for $k = \ell$ the M entries on the main diagonal are equal to 1 while all the remaining terms again vanish in norm as $1/N_R$. A similar statement also applies to the matrix $\mathbf{F}_{k,\ell,M}^t$, of course with entries vanishing as $1/N_T$. Accordingly, the following asymptotic formulas can be proven.

[11] In order to avoid an heavy notation, it is here dropped the dependence of the matrices $\mathbf{A}_{k,r}$ and $\mathbf{A}_{k,t}$ on the propagation paths arrival and departure angles, respectively.

(a) $N_T \to +\infty$, **finite** N_R: in this case the system becomes interference-free and it is obtained

$$\text{ASE} \approx \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T \gamma_k^2}{KM\sigma_n^2} \left(\mathbf{D}_k^H \mathbf{D}_k \right)^{-1} \mathbf{F}_{k,k,M}^{r\,H} \mathbf{L}_k \mathbf{L}_k^* \mathbf{F}_{k,k,M}^r \right| . \tag{4.37}$$

(b) $N_R \to +\infty$, **finite** N_T: it holds now

$$\text{ASE} \approx \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_M + \frac{P_T \gamma_k^2}{KM} \mathbf{R}_{D,\bar{k}}^{-1} \left(\mathbf{L}_k \mathbf{F}_{k,k,M}^{t\,H} \mathbf{F}_{k,k,M}^t \mathbf{L}_k^* \right)_{(1:M,1:M)} \right| , \tag{4.38}$$

with

$$\mathbf{R}_{D,\bar{k}} = \sigma_n^2 \mathbf{I}_M + \frac{P_T \gamma_k^2}{MK} \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \left(\mathbf{L}_k \mathbf{F}_{\ell,k,M}^{t\,H} \mathbf{F}_{\ell,k,M}^t \mathbf{L}_k^* \right)_{(1:M,1:M)} . \tag{4.39}$$

(c) $N_R, N_T \to \infty$: finally it holds

$$\text{ASE} \approx \sum_{k=1}^{K} \sum_{\ell=1}^{M} \log_2 \left(1 + \frac{P_T \gamma_k^2 |\alpha_{k,\ell}|^2}{KM\sigma_n^2} \right) . \tag{4.40}$$

It is easily seen that the above expression coincides with Eq. (4.28). Additionally, substituting Eqs. (4.37), (4.38) and (4.40) into the GEE definition (4.9), asymptotic expressions can be readily obtain for the system GEE, and, again, these can be maximized with respect to P_T by using Dinkelbach's algorithm.

Consider now the special case $M = 1$; the pre-coding and post-coding matrices are actually column vectors, and are expressed as

$$\mathbf{D}_k = \mathbf{a}_r(\phi_{1,k}^r) , \qquad \mathbf{Q}_k = \mathbf{a}_t(\phi_{1,k}^t) , \quad \forall k . \tag{4.41}$$

Using the above expressions, it is readily seen that Eq. (4.5) becomes

$$\text{ASE} = \sum_{k=1}^{K} \log_2 \left| 1 + \frac{P_T}{M} \mathbf{R}_{D,\bar{k}}^{-1} \left| \mathbf{a}_r^H(\phi_{1,k}^r) \mathbf{H}_k \mathbf{a}_t(\phi_{1,k}^t) \right|^2 \right| . \tag{4.42}$$

The interference covariance matrix $\mathbf{R}_{D,\bar{k}}$ is now just a scalar, and is written as

$$\mathbf{R}_{D,\bar{k}} = \sigma_n^2 + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \frac{P_T}{KM} \gamma_k^2 \left| \alpha_{k,1} f_{N_T}(\phi_{1,k}^t, \phi_{1,\ell}^t) + \sum_{i=2}^{N} \alpha_{k,i} f_{N_R}(\phi_{1,k}^r, \phi_{i,k}^r) f_{N_T}(\phi_{i,k}^t, \phi_{1,\ell}^t) \right|^2 . \tag{4.43}$$

Substituting Eq. (4.43) into Eq. (4.42) provides

$$\text{ASE} = \sum_{k=1}^{K} \log_2 \left(1 + \frac{\frac{P_T}{KM}\gamma_k^2 \left| \alpha_{k,1} + \sum_{i=2}^{N} \alpha_{k,i} f_{N_R}(\phi_{1,k}^r, \phi_{i,k}^r) f_{N_T}(\phi_{i,k}^t, \phi_{1k}^t) \right|^2}{\sigma_n^2 + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \frac{P_T}{KM}\gamma_k^2 \left| \alpha_{k,1} f_{N_T}(\phi_{1,k}^t, \phi_{1,\ell}^t) + \sum_{i=2}^{N} \alpha_{k,i} f_{N_R}(\phi_{1,k}^r, \phi_{i,k}^r) f_{N_T}(\phi_{i,k}^t, \phi_{1,\ell}^t) \right|^2} \right). \tag{4.44}$$

Eq. (4.44) provides the exact downlink ASE expression for finite values of N_T and N_R in the case of analog pre-coding and decoding, as a function of the reflection coefficients $\alpha_{\cdot,\cdot}$ and of the departure and arrival angles. In order to study its asymptotic values for large N_R and N_T, recall that $\gamma_k^2 = N_R N_T / N$.

(d) $N_T \to +\infty$, **finite** N_R: in this case the following holds

$$\text{ASE} \approx \sum_{k=1}^{K} \log_2 \left(1 + \frac{P_T}{KM} \frac{|\alpha_{k,1}|^2 N_T N_R}{N \sigma_n^2} \right). \tag{4.45}$$

It is seen that the ASE grows linearly with the number of users, logarithmically with the product $N_T N_R$, and the system is asymptotically interference-free and noise-limited. It can be also verified that the limiting ASE in Eq. (4.45) tends to the limiting ASE reported in Eq. (4.29), which holds for the case of FD beamforming, thus confirming the optimality of the considered analog beamforming in the limit of large number of transmit antennas.

(e) $N_R \to +\infty$, **finite** N_T: in this case the following holds

$$\begin{aligned} \text{ASE} &\approx \sum_{k=1}^{K} \log_2 \left(1 + \frac{\frac{P_T}{KM} \frac{N_R N_T}{N} |\alpha_{k,1}|^2}{\sigma_n^2 + \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \frac{P_T}{KM} \frac{N_R N_T}{N} \left| f_{N_T}(\phi_{1,k}^t, \phi_{1,\ell}^t) \right|^2 |\alpha_{k,1}|^2} \right) \\ &\to \sum_{k=1}^{K} \log_2 \left(1 + \frac{1}{\sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \left| f_{N_T}(\phi_{1,k}^t, \phi_{1,\ell}^t) \right|^2} \right). \end{aligned} \tag{4.46}$$

The ASE converges towards an asymptote that is independent of the number of receive antennas, while the system is now noise-free and interference-limited. The ASE now increases logarithmically with N_T^2, and there is no longer a linear increase of the ASE in the number of users. In particular, since for large K the quantity $\sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \left| f_{N_T}(\phi_{1,k}^t, \phi_{1,\ell}^t) \right|^2$ converges to $(K -$

$1)\mathbb{E}\left[\left|f_{N_T}(\phi^t_{1,k},\phi^t_{1,\ell})\right|^2\right]$, it can be shown that

$$\lim_{N_R,K\to+\infty}\text{ASE}=\frac{(\ln 2)^{-1}}{\mathbb{E}\left[\left|f_{N_T}(\phi^t_{1,k},\phi^t_{1,\ell})\right|^2\right]}\,. \tag{4.47}$$

Note that the above limiting value increases with N_T^2, while, for large K, the ASE per user vanishes.

(f) $N_R, N_T \to \infty$: in this case the same results as in 4.5.1(d) hold.

4.5.2 AN beamforming, uplink

For the uplink scenario, using the notation previously introduced, the k-th user ASE can be shown to be expressed as

$$\text{ASE}_k=\log_2\left|\mathbf{I}_M+\frac{P_{T,k}}{M}\gamma_k^2\mathbf{R}^{-1}_{U,\bar{k}}\mathbf{F}^r_{k,k,M}\mathbf{L}_k\;\mathbf{F}^{t\,H}_{k,k,M}\mathbf{F}^t_{k,k,M}\mathbf{L}_k^*\mathbf{F}^{r\,H}_{k,k,M}\right|\,, \tag{4.48}$$

with

$$\mathbf{R}_{U,\bar{k}}=\sigma_n^2\mathbf{D}_k^H\mathbf{D}_k+\sum_{\substack{\ell=1\\\ell\neq k}}^{K}\frac{P_{T,\ell}}{M}\gamma_\ell^2\mathbf{F}^r_{k,\ell,M}\mathbf{L}_\ell\mathbf{F}^{t\,H}_{\ell,\ell,M}\mathbf{F}^t_{\ell,\ell,M}\mathbf{L}_\ell^*\mathbf{F}^{r\,H}_{k,\ell,M}\,. \tag{4.49}$$

Asymptotic approximations for the k-th user ASE are now provided.

(a) $N_R \to +\infty$, **finite** N_T: in this case the system becomes interference-free and the following holds:

$$\text{ASE}_k\approx\log_2\left|\mathbf{I}_M+\frac{P_{T,k}\gamma_k^2}{M\sigma_n^2}\left(\mathbf{L}_k\mathbf{F}^{t\,H}_{k,k,M}\mathbf{F}^t_{k,k,M}\mathbf{L}_k^*\right)_{(1:M,1:M)}\right|\,. \tag{4.50}$$

(b) $N_T \to +\infty$, **finite** N_R:

$$\text{ASE}_k\approx\log_2\left|\mathbf{I}_M+\frac{P_{T,k}\gamma_k^2}{M}\mathbf{R}^{-1}_{U,\bar{k}}\mathbf{F}^r_{k,k,M}\mathbf{L}_k\mathbf{L}_k^*\mathbf{F}^{r\,H}_{k,k,M}\right|\,, \tag{4.51}$$

with

$$\mathbf{R}_{U,\bar{k}}=\sigma_n^2(\mathbf{D}_k^H\mathbf{D}_k)+\sum_{\substack{\ell=1\\\ell\neq k}}^{K}\frac{P_{T,\ell}\gamma_\ell^2}{M}\mathbf{F}^r_{k,\ell,M}\mathbf{L}_\ell\mathbf{L}_\ell^*\mathbf{F}^{r\,H}_{k,\ell,M}\,. \tag{4.52}$$

(c) $N_R, N_T \to \infty$: finally, the following holds

$$\text{ASE}_k\approx\sum_{\ell=1}^{M}\log_2\left(1+\frac{P_{T,k}\gamma_k^2|\alpha_{k,\ell}|^2}{M\sigma_n^2}\right)\,. \tag{4.53}$$

The above equation can be seen to be equal to Eq. (4.31).

Similarly to the downlink, also for the uplink, the case $M = 1$ permits skipping the matrix notation and obtaining more insightful formulas. For $M = 1$ the pre-coding and post-coding vectors are still given by Eq. (4.41), and the k-th user ASE in Eq. (4.7), after some algebra, is written as

$$\mathrm{ASE}_k = \log_2\left(1+\frac{\frac{P_{T,k}}{M}\gamma_k^2\left|\alpha_{k,1}+\sum_{i=2}^{N}\alpha_{k,i}f_{N_R}(\phi_{1,k}^r,\phi_{i,k}^r)f_{N_T}(\phi_{i,k}^t,\phi_{1k}^t)\right|^2}{\sigma_n^2+\sum_{\substack{\ell=1\\\ell\neq k}}^{K}\frac{P_{T,\ell}}{M}\gamma_\ell^2\left|\alpha_{\ell,1}f_{N_R}(\phi_{1,k}^r,\phi_{1,\ell}^r)+\sum_{i=2}^{N}\alpha_{\ell,i}f_{N_T}(\phi_{i,\ell}^t,\phi_{1,\ell}^t)f_{N_R}(\phi_{1,k}^r,\phi_{i,\ell}^r)\right|^2}\right). \tag{4.54}$$

Eq. (4.54) provides the downlink ASE expression for finite values of N_T and N_R in the case of analog pre-coding and decoding, as a function of the reflection coefficients $\alpha_{\cdot,\cdot}$ and of the departure and arrival angles. In order to study its asymptotic values for large N_R and N_T, recall that $\gamma_k^2 = N_R N_T / N$.

(d) $N_R \to +\infty$, **finite** N_T: in this case the following expression holds

$$\mathrm{ASE}_k \to \log_2\left(1+\frac{P_{T,k}}{M}\frac{|\alpha_{k,1}|^2 N_T N_R}{N\sigma_n^2}\right). \tag{4.55}$$

It is seen that the total ASE grows linearly with the number of users, logarithmically with the product $N_T N_R$, and the system is asymptotically interference-free and noise-limited. It can be also verified that the limiting ASE in Eq. (4.55) tends to coincide with the limiting ASE reported in Eq. (4.29), which holds for the case of FD beamforming, thus confirming the optimality of the considered analog beamforming in the limit of large number of transmit antennas.

(e) $N_T \to +\infty$, **finite** N_R: in this case the following relation holds

$$\begin{aligned}\mathrm{ASE}_k &\approx \log_2\left(1+\frac{\frac{P_{T,k}}{M}\frac{N_R N_T}{N}|\alpha_{k,1}|^2}{\sigma_n^2+\sum_{\substack{\ell=1\\\ell\neq k}}^{K}\frac{P_{T,\ell}}{M}\frac{N_R N_T}{N}\left|f_{N_R}(\phi_{1,k}^r,\phi_{1,\ell}^r)\right|^2|\alpha_{\ell,1}|^2}\right)\\ &\to \log_2\left(1+\frac{|\alpha_{k,1}|^2}{\sum_{\substack{\ell=1\\\ell\neq k}}^{K}|\alpha_{\ell,1}|^2\left|f_{N_R}(\phi_{1,k}^r,\phi_{1,\ell}^r)\right|^2}\right).\end{aligned} \tag{4.56}$$

The ASE converges towards an asymptote that is independent of the number of transmit antennas, while the system is now noise-free and interference-limited. The ASE now increases logarithmically with N_R^2. For large K, the

following holds:

$$\sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} |\alpha_{\ell,1}|^2 \left| f_{N_R}(\phi^r_{1,k}, \phi^r_{1,\ell}) \right|^2 \approx (K-1)\mathbb{E}\left[|\alpha_{\ell,1}|^2 \left| f_{N_R}(\phi^r_{1,k}, \phi^r_{1,\ell}) \right|^2 \right] ,$$

and also in this case the ASE per user vanishes.

(f) $N_R, N_T \to \infty$: in this case the same results as in 4.5.2(d) hold.

4.6 Numerical results

Simulation results showing the ASE and the GEE for a single-cell mm-Wave MIMO system are now provided; it is assumed $K = 10$ users use the same frequency band with random locations at a maximum distance 100 m from the BS. The parameters for the generation of the matrix channels are reported in [32] for the "street canyon model", with $N_{\text{cl}} = 2$ and $N_{\text{ray}} = 20$. The carrier frequency is $f_c = 73$ GHz, the used bandwidth is $W = 500$ MHz[12], the noise power $\sigma_n^2 = F\mathcal{N}_0 W$, with receiver noise figure $F = 3$ dB and $\mathcal{N}_0 = -174$ dBm/Hz. All the considered low-complexity beamformers have been realized using a number of RF chains equal to the multiplexing order M (and KM at the BS). The results come from an average over 500 independent realizations of users' locations and propagation channels.

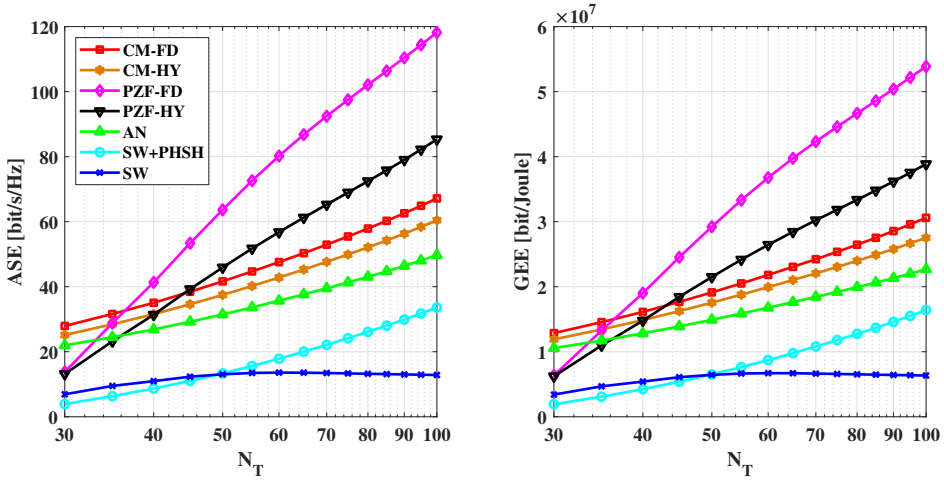

FIGURE 4.3: Plot of downlink ASE and GEE versus N_T with $N_R = 30$, $K = 10$, $M = 3$ and $P_T = 0$ dBW.

[12]Standardization bodies have not yet set the mm-Wave carrier frequencies that will be really used in practice. However, the considered values can be deemed as representative of a typical mm-Wave link for wireless cellular communications.

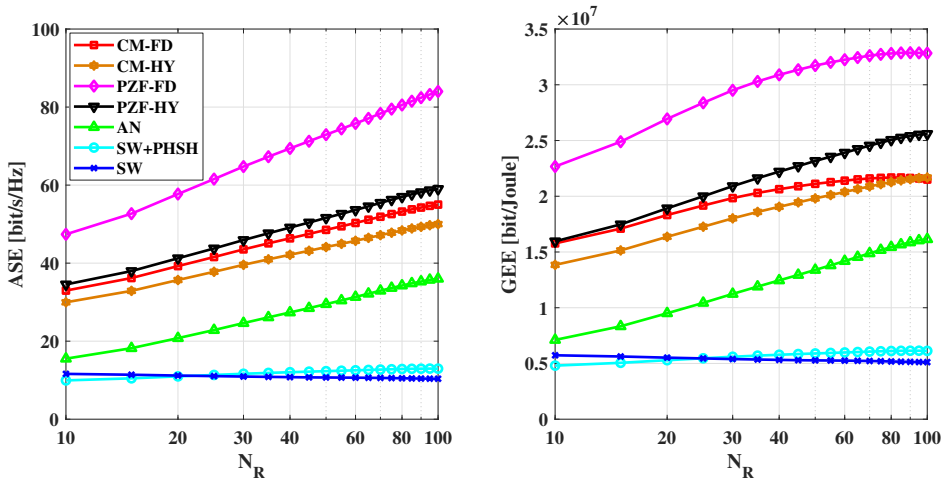

FIGURE 4.4: Plot of downlink ASE and GEE versus N_R with $N_T = 50$, $K = 10$, $M = 3$ and $P_T = 0$ dBW.

First of all, results as a function of the number of transmit and receive antennas are reported. Figs. 4.3 and 4.4 report the downlink ASE and the GEE versus the number of transmit antennas (assuming $N_R = 30$) and versus the number of receive antennas (assuming $N_T = 50$), respectively, assuming $P_T = 0$ dBW and multiplexing order $M = 3$. Results corresponding to all the previously detailed beamforming structures are reported. Inspecting the figures, it is seen that the best performing beamforming structure is the PZF-FD, both in terms of ASE and of GEE[13]. This last conclusion is quite surprising, since it shows that lower complexity structures, although necessary for obvious practical considerations, actually are less energy efficient (from a communication physical layer perspective) than FD structures. Results also show that the SW structure achieves quite unsatisfactory performance; moreover, for low values of N_T the CM-FD and its HY approximation outperform the PZF-FD and PZF-HY solutions. From Fig. 4.4 it can be also seen that while the ASE grows with the number of antennas, the GEE instead exhibits a maximum: in particular, it is seen that, for the considered scenario, the PZF-FD beamformer achieves its maximum GEE for $N_T \approx 90$.

[13]Note however that for small values of N_T the PZF beamforming structures achieve inferior performance with respect to the other solutions due to the reduced dimensionality of the interference-free subspace.

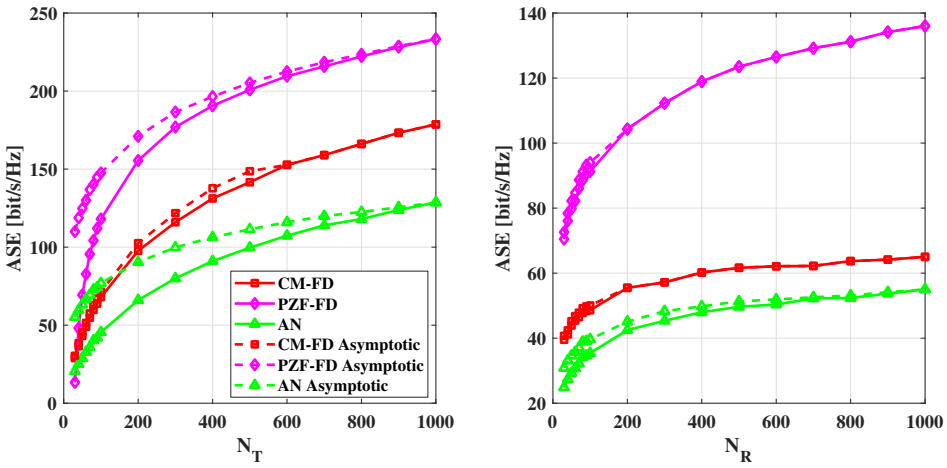

FIGURE 4.5: Plot of asymptotic downlink ASE formulas versus N_T and N_R, with $K = 10$, $M = 3$ and $P_T = 0$ dBW.

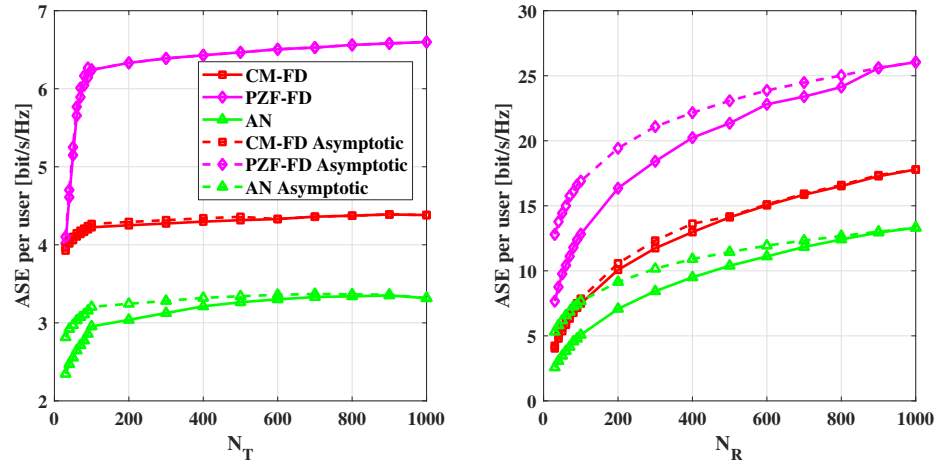

FIGURE 4.6: Plot of asymptotic uplink ASE per user formulas versus N_T and N_R, with $K = 10$, $M = 3$ and $P_T = 0$ dBW.

Figs. 4.5 and 4.6 are devoted to the validation of the derived asymptotic formulas in the large number of antennas regime. In particular, the subplots in Fig. 4.5 show the downlink ASE, versus N_T (assuming $N_R = 30$) and versus N_R (assuming $N_T = 50$), for the CM-FD, PZF-FD and AN beamformers, and their asymptotic approximation reported in Eqs. (4.25), (4.29), (4.37), and (4.38). The subplots in Fig. 4.6, instead, refer to the uplink and report the ASE per user, again versus N_T (assuming $N_R = 30$) and versus N_R (assuming $N_T = 50$), for the CM-FD, PZF-FD and AN beamformers, and their asymptotic approximation reported in Eqs. (4.27), (4.31), (4.50), and (4.51). Results fully confirm the effectiveness of the found asymptotic formulas, that may turn out to be useful in the derivation of simplifies resource allocation strategies.

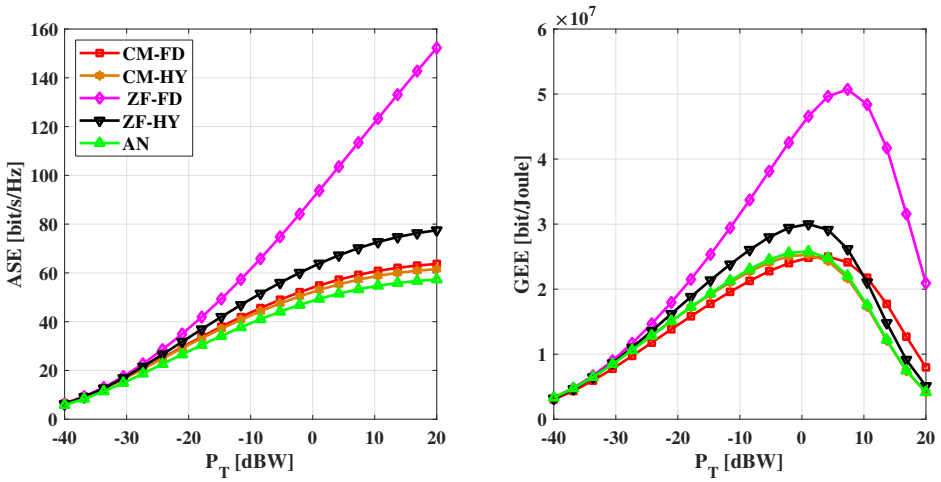

FIGURE 4.7: Plot of downlink ASE and GEE versus P_T for a system with $N_T = 100$, $N_R = 30$, $K = 10$ and $M = 1$.

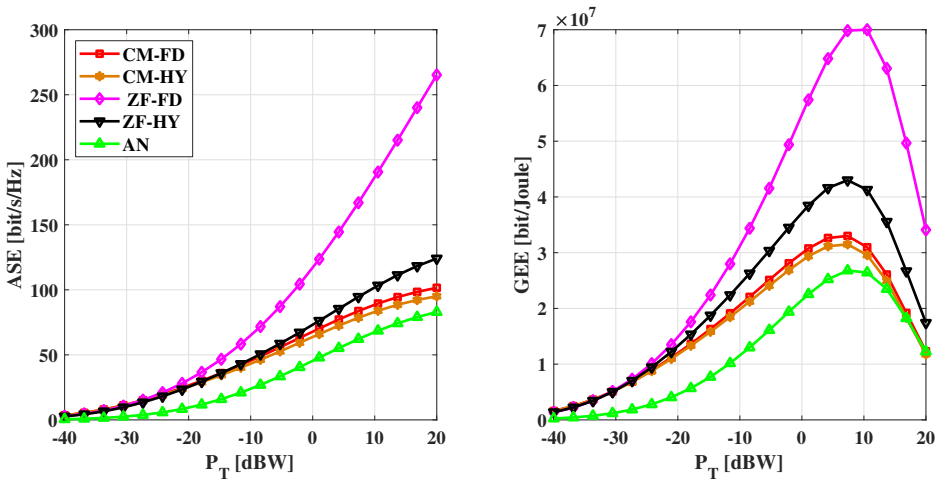

FIGURE 4.8: Plot of downlink ASE and GEE versus P_T for a system with $N_T = 100$, $N_R = 30$, $K = 10$ and $M = 3$.

Figs. 4.7 and 4.8, finally, report the downlink system ASE and GEE versus the transmit power, for the case of multiplexing order $M = 1$ and $M = 3$. Here, a system with $N_T = 100$ antennas at the BS and $N_R = 30$ antennas at the MSs has been considered. Also in this case the number of users is $K = 10$. Results show a trend that has already been found elsewhere (e.g., in [76]); in particular, while the ASE grows with the transmit power (at least in the considered range of values), the GEE exhibits instead a maximum around 0 dBW. This behavior is explained by the fact that for large values of the transmit power, the numerator in the GEE grows at a slower rate than the denominator of the GEE, and so the GEE itself decreases. From an energy-efficiency perspective, increasing the transmit power beyond the GEE-optimal point leads to moderate improvements in the system throughput at the price of a much higher increase in the consumed power.

Chapter 5

User-centric approach to cell-free massive MIMO networks

Referred articles published in the *IEEE Wireless Communications Letters* 2017, in the *Proc. of WSA 2017; 21th ITG International Workshop on Smart Antennas* and in the *Proc. of 2018 15th International Symposium on Wireless Communication Systems (ISWCS)*.

5.1 Introduction

As discussed in Chapter 1, massive MIMO is a promising wireless access technology that can provide high throughput with simple signal processing for 5G and beyond-5G networks [1]. Massive antenna at the BSs can be deployed in co-located or distributed setups. In co-located massive MIMO all the antennas are located in a compact area and this architecture has the advantage of low backhaul requirements. In distributed massive MIMO systems, instead, the antennas are spread out over a large area; this architecture has the advantage of efficiently exploiting mascroscopic diversity against the shadow fading, so these systems can potentially offer a much higher probability of coverage than collocated massive MIMO [77], at the cost of increased backhaul requirements. Additionally, the distributed layout permits alleviating the cell-edge problem, since it considerably lowers the probability that a user is situated far from every system AP, permitting thus to achieve a better fairness and service uniformity across mobile users. In [78], the viability of using distributed antennas in multi-cell systems for massive MIMO on the uplink is investigated for a particular spatial correlation channel model. In [79] the authors focus on the downlink of a multicell distributed antenna system assuming that only the slowly-varying large-scale channel state is required at the transmitter and they explore the performance gain that can be achieved by coordinated transmissions for a virtual MIMO system. One of the drawbacks of such virtual MIMO systems is the heavy backhaul requirements, since, besides data symbols, also the channel estimates and

the beamforming schemes are to be shared with the CPU. Recently, a CF massive MIMO architecture has been introduced where a very large number of distributed single-antenna APs serve many single-antenna MSs in the same time-frequency resource [35], [80]. All APs are connected to a CPU and cooperate via a backhaul network, serving all the MSs via TDD operation, so that in a CF massive MIMO system there are actually no cells or cell boundaries. The CF concept thus can be interpreted as a scalable and lower-complexity implementation of distributed massive MIMO or of ultra-dense AP deployments with cloud-based radio access networks.
The distinguishing features of the CF architecture are the following:

(a) the time division duplex protocol is used to exploit channel reciprocity on the uplink and downlink;

(b) uplink channel estimates are computed locally at each AP and exploited locally, which means that they are not sent on the backhaul link;

(c) beamformers to be used at the APs are computed locally and not at the CPU;

(d) the backhaul is used to send data symbols on the downlink and sufficient statistics on the uplink to perform centralized uplink data decisions; however, it is not used to share the channel estimates and/or the beamformers.

CF massive MIMO systems have received increasing attention in the recent past. The authors of [35] show that the CF approach provides better performance than a small-cell system in terms of 95%-likely per-user throughput, thus confirming that the scheme is effective in alleviating the cell-edge user problem and in providing a more uniform service across users. CF massive MIMO is a recent research topic that however has been gaining a strong momentum in the last few years. The paper [81] has shown that some performance improvement can be obtained in low density networks by using downlink pilots, while the paper [82], instead, analyzes the performance improvements granted by the use of a zero-forcing precoder in the downlink: although the gains are from five to ten-fold, the zero-forcing precoder requires centralized computations at the CPU and increased backhaul overhead. Zero-forcing pre-coding is again considered in [83], wherein it is coupled with a power control algorithm aimed at maximizing the energy efficiency of CF massive MIMO considering the backhaul power consumption and the imperfect CSI. In [84], the uplink performance of CF systems is investigated using MMSE processing for the case in which the energy efficiency of CF massive MIMO is to be maximized considering the backhaul power consumption and the imperfect CSI. The energy-efficiency of CF massive MIMO systems is also considered in [85], which proposes a power allocation algorithm aiming at maximizing the total energy efficiency, subject to a per-user

spectral efficiency constraint and a per-AP power constraint. The power allocation strategies here are simplified by the fact that single-antenna transceivers are considered both at the APs and at the MSs, which permit skipping $\log|(\cdot)|$ functions in the achievable rate formulas. In order to reduce the backhaul overhead, the paper [86] considers instead a coded CF massive MIMO system, and investigates the performance of a compute & forward mechanism for the uplink, wherein each AP attempts to use an integer linear combination of the codewords to represent the scaled received signal to be sent to the CPU. The finite backhaul capacity is also considered in [87], which studies the case in which quantized version of the estimated channel and the quantized received signal are available at the CPU, and the case when only the quantized version of the combined signal with MR post-coding detector is available at the CPU. It should be noted that all the cited papers consider the case in which both the APs and the MSs are equipped with a single-antenna, the only exception being references [85], [87], which consider multiple antennas at the APs. The extension of the CF massive MIMO architecture to the case in which also the MSs are equipped with multiple antennas is not trivial since no channel estimation is performed at the MSs, and so no channel-dependent beamforming scheme can be used there.

One critical point of the originally formulated version of the CF massive MIMO systems [35], [80] is the fact that all the APs serve all the MSs in the system. This assumption may lead to some inefficiencies in the system as the size of the considered area grows: indeed, it appears clearly pointless to waste power and computational resources at an AP to decode MSs that are very far and that are presumably received with a very low SINR. To overcome this limitation, a UC distributed massive MIMO system has been introduced, still for single-antenna APs and MSs, in [88]; in the UC approach, each MS is served not by all the APs in the system, but just by the ones that are in the neighborhood. Similar APs selection strategies, a received-power-pased selection and a largest-large-scale-fading-based selection, exploiting a partial knowledge of the channel statistic, are proposed in reference [85]. The UC approach, while being much simpler than the CF one and less hungry of backhaul bandwidth, was shown in [88] to provide a larger achievable rate-per-user to the majority of the MSs in the system.

Following on such a track, and building upon the conference papers [89], [90], this chapter of the thesis provides a thorough comparison of the UC and CF approaches, considering the case in which the MSs and the APs are equipped with multiple antennas. In the UC approach, each AP communicates only with a pre-assigned number of MSs, to which it has best link conditionos. The overall complexity of the system is thus lower than CF, since, every AP has to serve only a pre-assigned number of MSs, and, moreover, a reduced amount of data is to be sent on the backhaul links to the CPU. In this chapter of the thesis a

beamforming scheme that does not require channel estimation at the MSs is introduced; rather the proposed scheme exploits the channel hardening effect due to the large number of antennas in order to perform coherent data reception at the MSs. Moreover, a simple and low complexity pilot matched (PM) channel estimation strategy implemented at each APs is proposed and maximum-length-sequences (pseudo-noise) are assumed as pilots at each MS in the uplink training phase. Channel inversion beamforming is proposed here as a generalization of the conjugate beamforming applied in the single-antenna case, and, again, as the in the originally formulated version of the CF massive MIMO, no channel estimates and beamforming matrices are propagated through the backhaul network. Furthermore, in this chapter of the thesis two power allocation strategies for the uplink and the downlink are proposed, both for the CF and the UC case. The power allocation policy follows the APs selection in the UC case, i.e., only the power of the APs involved in the UC communication is allocated. The first power allocation strategy is a sum-rate maximizing power allocation one, aimed at maximizing performance of the system in terms of overall data-rate and the second strategy is a minimum-rate maximizing power allocation one, aimed at maximizing performance of the system in terms of fairness. Both the optimization problems have non-concave objective functions, which makes their solution challenging. In the proposed power allocation strategies it is here used the *successive lower-bound maximization*, that merges the tools of *alternating optimization* and of the *sequential convex programming*, in order to solve the two optimization problems and to obtain the power allocation algorithms detailed in the chapter of the thesis. Additionally, advanced signal processing techniques for uplink channel estimation, uplink data detection and downlink pre-coding are proposed in this chapter. In particular, with regard to channel estimation, a linear MMSE channel estimation procedure is detailed. For uplink data detection, a local SIC scheme is proposed, in the sense that partial interference cancellation is implemented at the APs by using locally available information only, so that in the sufficient statistics sent to the CPU the interference has already been pruned without sending all the channel estimates at the CPU. Finally, with regard to downlink data transmission, a local PZF beamforming architecture is introduced, where the beamformer is computed at each AP and synthesized by using locally available information obtained during the training phase, and exploiting the availability of multiple antennas at the AP. The results obtained with the two power allocation strategies proposed in the chapter of the thesis are contrasted with the case of uniform power allocation, i.e., all the APs and the MSs transmit data with the maximum power available, and the CF and UC approaches with PM channel estimation and with perfect CSI are condered. Results will show that the UC approach generally outperforms the CF one, especially on the uplink. Regarding the advanced signal processing techniques, numerical results show the effectiveness of the proposed solutions; in particular

the use of SIC on the uplink can more than double the median rate in a typical setting compared to the simple MR post-coding in the uplink.
The remainder of this chapter is organized as follows. Next Section contains the description of the considered system model. Section 5.3 is devoted to the illustration of the comunication protocol, composed by uplink training, downlink data transmission and uplink data transmission, for both CF and UC approaches. In Section 5.4 the performance measures and the two power control strategies proposed for the downlink are reported. Section 5.5 contains the two power control strategies proposed for the uplink. Section 5.6 details the proposed advanced signal processing techniques and Section 5.7 contains the numerical results.

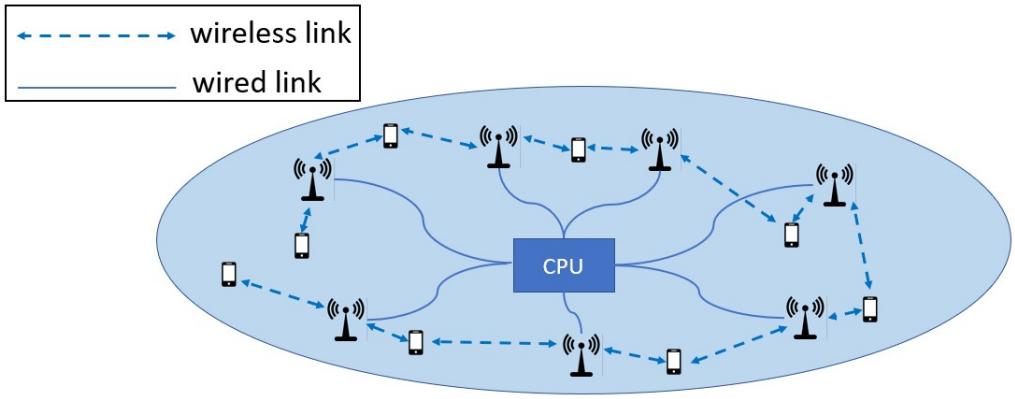

FIGURE 5.1: UC approach to CF massive MIMO system.

5.2 System model

In this chapter an area with K MSs and M APs is considered (see Fig. 5.1), where MSs and APs are randomly located. The M APs are connected by means of a backhaul network to a CPU wherein data-decoding is performed. In accordance to the approach of [35], [80], all communications take place on the same frequency band; uplink and downlink are separated through TDD; the coherence interval is thus divided into three phases: (a) uplink channel estimation, (b) downlink data transmission, and (c) uplink data transmission. In phase (a) the MSs send pilot data in order to enable channel estimation at the APs. In phase (b) APs use channel estimates to perform CM beamforming and send data symbols on the downlink; while in the CF architecture APs send data to all the MSs in the system, in the UC approach APs send data only to a subset of the MSs in the system. Finally, in phase (c) MSs send uplink data symbols to the APs; while in the CF architecture all the APs participate to the decoding of the data transmitted by all the MSs, in the UC approach APs just decode the data from the nearby MSs. The procedure for the selection of the MSs to serve will be specified in the following section. No pilots are transmitted on the downlink and no

channel estimation is performed at the MSs: data decoding takes place on the downlink relying on the fact that in TDD the downlink channel is the reciprocal of the uplink channel[14] and on the channel hardening effect due to many transmitting APs. In the following, N_{MS} and N_{AP} denote the number of antennas at the MSs and at the APs, respectively.

- **Channel model** The $(N_{\text{AP}} \times N_{\text{MS}})$-dimensional matrix $\mathbf{G}_{k,m}$ denotes the channel between the k-th MS and the m-th AP:

$$\mathbf{G}_{k,m} = \sqrt{\beta_{k,m}}\mathbf{H}_{k,m}\,, \tag{5.1}$$

with $\beta_{k,m}$ a scalar coefficient modeling the channel shadowing effects and $\mathbf{H}_{k,m}$ an $(N_{\text{AP}} \times N_{\text{MS}})$-dimensional matrix whose entries are i.i.d $\mathcal{CN}(0,1)$ random variables (RVs). For the path loss and the shadow fading correlation models, the ones reported in [35] are used in this thesis. The large scale coefficient $\beta_{k,m}$ in Eq. (5.1) models the path loss and shadow fading, according to Eq. (3.2), where the path loss PL and the shadow fading coefficient z between the m-th AP and the k-th user are replaced with $\text{PL}_{k,m}$ and $z_{k,m}$, respectively. For the path loss, the three slope path loss model in Eq. (3.3) is used, where d is replaced with $d_{k,m}$, that denotes the distance between the m-th AP and the k-th user (expressed in km). In real-world scenarios, transmitters and receivers that are in close vicinity of each other may be surrounded by common obstacles, and hence, the shadow fading RVs are correlated; for the shadow fading coefficient it is thus used a model with two components [92]

$$z_{k,m} = \sqrt{\delta}a_m + \sqrt{1-\delta}b_k, \quad m = 1,\ldots,M,\ k = 1,\ldots,K, \tag{5.2}$$

where $a_m \sim \mathcal{N}(0,1)$ and $b_k \sim \mathcal{N}(0,1)$ are independent RVs, and $\delta,\ 0 \leq \delta \leq 1$ is a parameter. The covariance functions of a_m and b_k are given by:

$$\mathbb{E}\left[a_m a_{m'}\right] = 2^{-\frac{d_{\text{AP(m,m')}}}{d_{\text{decorr}}}} \qquad \mathbb{E}\left[b_k b_{k'}\right] = 2^{-\frac{d_{\text{MS(k,k')}}}{d_{\text{decorr}}}}\,, \tag{5.3}$$

where $d_{\text{AP(m,m')}}$ is the geographical distance between the m-th and m'-th APs, $d_{\text{MS(k,k')}}$ is the geographical distance between the k-th and the k'-th MSs. The parameter d_{decorr} is a decorrelation distance which depends on the environment, typically this value is in the range 20-200 m.

5.3 Communication protocol for the CF and UC approaches

As already discussed, the communication procedure is made of three different phases, (a) uplink training, (b) downlink data transmission, and (c) uplink data

[14] According to [35], the channel reciprocity is also ensured by perfect hardware chain calibration, whose feasibility has been recently shown in [91].

transmission. The overall duration of these three phases must not exceed the channel coherence time, thus implying that these three phases must be sequentially repeated with a frequency larger than the channel Doppler spread.

5.3.1 Uplink training

During this phase the MSs send uplink training pilots in order to permit channel estimation at the APs. This phase is the same for both the UC and CF approaches. The length (in samples) of the channel coherence time is denoted by τ_c , and the length (in samples) of the uplink training phase is denoted by τ_p, of course $\tau_p < \tau_c$. Denote by $\mathbf{\Phi}_k \in \mathcal{C}^{N_{\mathrm{MS}} \times \tau_p}$ the pilot sequence sent by the k-th MS, and assume that the rows of $\mathbf{\Phi}_k$ have unit norm. The signal received at the m-th AP in the n-th signaling time is represented by the following N_{AP}-dimensional vector

$$\mathbf{y}_m(n) = \sum_{k=1}^{K} \sqrt{p_k} \mathbf{G}_{k,m} \mathbf{\Phi}_k(:, n) + \mathbf{w}_m(n) \ , \tag{5.4}$$

with $\sqrt{p_k}$ the user k transmit power on each antenna during the training phase. Collecting all the observable vectors $\mathbf{y}_m(n)$, for $n = 1, \ldots, \tau_p$ into the $(N_{\mathrm{AP}} \times \tau_p)$-dimensional matrix $\mathbf{Y}_m$, it is easy to show that

$$\mathbf{Y}_m = \sum_{k=1}^{K} \sqrt{p_k} \mathbf{G}_{k,m} \mathbf{\Phi}_k + \mathbf{W}_m \ . \tag{5.5}$$

In the above equation the matrix $\mathbf{W}_m$ is $(N_{\mathrm{AP}} \times \tau_p)$-dimensional and contains the thermal noise contribution and out-of-cell interference at the m-th AP; its entries are assumed to be i.i.d. $\mathcal{CN}(0, \sigma_w^2)$ random variables. Based on the observable matrix $\mathbf{Y}_m$, the m-th AP performs estimation of the channel matrices $\{\mathbf{G}_{k,m}\}_{k=1}^{K}$. Simple PM single-user channel estimation is here assumed, for the sake of simplicity (more sophisticated channel estimation schemes, based, e.g., on linear minimum mean square error processing, might however be considered) and the knowledge of MSs transmit powers $\{p_k\}_{k=1}^{K}$ is assumed. The estimate, $\widehat{\mathbf{G}}_{k,m}$ say, of the channel matrix $\mathbf{G}_{k,m}$ is

$$\widehat{\mathbf{G}}_{k,m} = \frac{1}{\sqrt{p_k}} \mathbf{Y}_m \mathbf{\Phi}_k^H = \mathbf{G}_{k,m} \mathbf{\Phi}_k \mathbf{\Phi}_k^H + \sum_{\substack{j=1 \\ j \neq k}}^{K} \sqrt{\frac{p_j}{p_k}} \mathbf{G}_{j,m} \mathbf{\Phi}_j \mathbf{\Phi}_k^H + \frac{1}{\sqrt{p_k}} \mathbf{W}_m \mathbf{\Phi}_k^H \ . \tag{5.6}$$

Estimation (5.6) must be made in all the APs (i.e., for all the values of $m = 1, \ldots, M$) for all the values of $k = 1, \ldots, K$. If the rows of the matrices $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_K$ are pairwisely orthogonal (i.e., $\mathbf{\Phi}_k \mathbf{\Phi}_j = \mathbf{I}_{N_{\mathrm{MS}}} \delta_{i,k}$, for all i, k), then Eq. (5.6) simplifies to

$$\widehat{\mathbf{G}}_{k,m} = \frac{1}{\sqrt{p_k}} \mathbf{Y}_m \mathbf{\Phi}_k^H = \mathbf{G}_{k,m} + \frac{1}{\sqrt{p_k}} \mathbf{W}_m \mathbf{\Phi}_k^H \ , \tag{5.7}$$

and thermal noise is the only disturbance impairing the channel estimate. A necessary condition for this to happen is however $\tau_p \geq K N_{\mathrm{MS}}$, a relation that

usually is not verified in practical scenarios due to the fact that τ_p must be a fraction of the channel coherence length. As a consequence, almost orthogonal pilot sequences are usually employed. The pilot sequences assigned to each user are assumed mutually orthogonal, so that $\mathbf{\Phi}_k \mathbf{\Phi}_k^H = \mathbf{I}_{N_{\mathrm{MS}}}$, while, instead, pilot sequences from different users are non-orthogonal. As a consequence, Eq. (5.6) is actually expressed as:

$$\widehat{\mathbf{G}}_{k,m} = \mathbf{G}_{k,m} + \sum_{\substack{j=1 \\ j \neq k}}^{K} \sqrt{\frac{p_j}{p_k}} \mathbf{G}_{j,m} \mathbf{\Phi}_j \mathbf{\Phi}_k^H + \frac{1}{\sqrt{p_k}} \mathbf{W}_m \mathbf{\Phi}_k^H \,, \tag{5.8}$$

which clearly shows that the channel estimate is degraded not only by noise, but also by the pilots from the other users, i.e., the pilot contamination phenomenon.

5.3.2 Downlink data transmission

After each AP has obtained estimates of the channel matrix from all the MSs in the system, the downlink data transmission phase begins. The APs treat the channel estimates as the true channels, and channel inversion beamforming is performed to transmit data to the MSs. The objective of this beamforming scheme is to ensure that the MSs will be able to receive data without information on the channel state. Denote by P_k the multiplexing order (i.e., the number of simultaneous data-streams) for user k, and by $\mathbf{x}_k^{\mathrm{DL}}(n)$ the P_k-dimensional unit-norm vector containing the k-th user data symbols to be sent in the n-th sample time, let $\mathbf{L}_k = \mathbf{I}_{P_k} \otimes \mathbf{1}_{N_{\mathrm{MS}}/P_k}$ be the channel independent beamformer to be used at each MS. Basically, this corresponds to partitioning the MS antennas in as many disjoint subsets as the multiplexing order, and to use all the antennas in the same subset to transmit and receive one data stream. Note that this channel independent beamforming at the MSs, based on its simple definition, remains the same in the two CF and UC approaches. The downlink precoder at the m-th AP for the k-th MS is expressed as

$$\mathbf{Q}_{k,m} = \widehat{\mathbf{G}}_{k,m} \left(\widehat{\mathbf{G}}_{k,m}^H \widehat{\mathbf{G}}_{k,m} \right)^{-1} \mathbf{L}_k \,. \tag{5.9}$$

It is easy to realize that using the above precoder at the AP, coupled with the use of the channel independent beamformer $\mathbf{L}_k$ at the MS, permits inverting the channel effect and perfectly recovering the transmitted symbols, in the ideal case of perfect channel knowledge, $\mathbf{L}_k^H \mathbf{G}_{k,m}^H \mathbf{Q}_{k,m} = \mathbf{I}_{P_k}$. Note also that computing the beamformers in Eq. (5.9) requires a computational effort proportional to $N_{\mathrm{AP}} N_{\mathrm{MS}}^3$, which is easily manageable since in the considered distributed environment, and for the case of sub-6 GHz frequencies, both N_{AP} and N_{MS} are small numbers.

- **CF massive MIMO architecture**
  In the CF architecture all the APs communicate with all the MSs in the

systems, so the signal transmitted by the m-th AP in the n-th interval is the following N_{AP}-dimensional vector

$$\mathbf{s}_m^{\mathrm{cf}}(n) = \sum_{k=1}^{K} \sqrt{\eta_{k,m}^{\mathrm{DL,cf}}} \mathbf{Q}_{k,m} \mathbf{x}_k^{\mathrm{DL}}(n) \ , \tag{5.10}$$

with $\eta_{k,m}^{\mathrm{DL,cf}}$ a scalar coefficient ruling the power transmitted by the m-th AP for the k-th MS. The generic k-th MS receives signal contributions from all the APs; the observable vector is expressed as

$$\begin{aligned} \mathbf{r}_k^{\mathrm{cf}}(n) &= \sum_{m=1}^{M} \mathbf{G}_{k,m}^H \mathbf{s}_m^{\mathrm{cf}}(n) + \mathbf{z}_k(n) \\ &= \sum_{m=1}^{M} \sqrt{\eta_{k,m}^{\mathrm{DL,cf}}} \mathbf{G}_{k,m}^H \mathbf{Q}_{k,m} \mathbf{x}_k^{\mathrm{DL}}(n) \\ &+ \sum_{m=1}^{M} \sum_{\substack{j=1 \\ j \neq k}}^{K} \sqrt{\eta_{j,m}^{\mathrm{DL,cf}}} \mathbf{G}_{k,m}^H \mathbf{Q}_{j,m} \mathbf{x}_j^{\mathrm{DL}}(n) + \mathbf{z}_k(n) \ . \end{aligned} \tag{5.11}$$

In Eq. (5.11), the N_{MS}-dimensional vector $\mathbf{z}_k(n)$, modelled as i.i.d. $\mathcal{CN}(0, \sigma_z^2)$ random variables, represents the thermal noise and out-of-cluster interference at the k-th MS. Based on the observation of the vector $\mathbf{r}_k^{\mathrm{cf}}(n)$, a soft estimate of the data symbols $\mathbf{x}_k^{\mathrm{DL}}(n)$ is obtained at the k-th MS as

$$\widehat{\mathbf{x}}_k^{\mathrm{DL,cf}}(n) = \mathbf{L}_k^H \mathbf{r}_k^{\mathrm{cf}}(n) \ . \tag{5.12}$$

Note once again that no channel estimation is performed at the MSs; the beamformers $\mathbf{L}_k$ have a fixed structure independent of the channel realization, so that the entries of the observation vector are partitioned in P_k groups and a coherent sum is made within each group.

- **UC massive MIMO architecture**
  In the UC approach, the APs communicate only with the closest MSs. In order to define a measure for the closeness of the MSs, several procedures can be conceived. One possible strategy is that each AP computes the average Frobenius norm of the estimated channels for all the MSs, i.e.,

$$\bar{\mathbf{G}}_m = \frac{1}{K} \sum_{k=1}^{K} \|\widehat{\mathbf{G}}_{k,m}\|_F, \tag{5.13}$$

  and will serve only the APs whose channel estimates have a Frobenius norm larger than the computed average value. Another possible approach is that each AP sorts these estimates in descending Frobenius norm order and serves only the N MSs with the strongest channel, with N a proper design parameter. In this chapter of the thesis numerical results using this latter strategy are presented. The set of MSs served by the m-th AP is

denoted by $\mathcal{K}(m)$. Given the sets $\mathcal{K}(m)$, for all $m = 1, \ldots, M$, it is possible to define the set $\mathcal{M}(k)$ of the APs that communicate with the k-th user:

$$\mathcal{M}(k) = \{m : k \in \mathcal{K}(m)\}. \tag{5.14}$$

So, in this case, the signal transmitted by the m-th AP in the n-th interval is the N_{AP}-dimensional vector

$$\mathbf{s}_m^{\mathrm{uc}}(n) = \sum_{k \in \mathcal{K}(m)} \sqrt{\eta_{k,m}^{\mathrm{DL,uc}}} \mathbf{Q}_{k,m} \mathbf{x}_k^{\mathrm{DL}}(n) \ , \tag{5.15}$$

with $\eta_{k,m}^{\mathrm{DL,uc}}$, again, a scalar coefficient ruling the power transmitted by the m-th AP. The generic k-th MS receives signal contributions from all the APs; the observable vector is expressed as

$$\begin{aligned} \mathbf{r}_k^{\mathrm{uc}}(n) &= \sum_{m=1}^{M} \mathbf{G}_{k,m}^H \mathbf{s}_m^{\mathrm{uc}}(n) + \mathbf{z}_k(n) \\ &= \sum_{m \in \mathcal{M}(k)} \sqrt{\eta_{k,m}^{\mathrm{DL,uc}}} \mathbf{G}_{k,m}^H \mathbf{Q}_{k,m} \mathbf{x}_k^{\mathrm{DL}}(n) \\ &+ \sum_{\substack{j=1 \\ j \neq k}}^{K} \sum_{m \in \mathcal{M}(j)} \sqrt{\eta_{j,m}^{\mathrm{DL,uc}}} \mathbf{G}_{k,m}^H \mathbf{Q}_{j,m} \mathbf{x}_j^{\mathrm{DL}}(n) + \mathbf{z}_k(n) \ . \end{aligned} \tag{5.16}$$

In Eq. (5.16), the N_{MS}-dimensional vector $\mathbf{z}_k(n)$ represents the thermal noise and out-of-cluster interference at the k-th MS, and is modeled as i.i.d. $\mathcal{CN}(0, \sigma_z^2)$ RVs. Based on the observation of the vector $\mathbf{r}_k^{\mathrm{uc}}(n)$, a soft estimate of the data symbols $\mathbf{x}_k^{\mathrm{DL}}(n)$ is obtained at the k-th MS as

$$\widehat{\mathbf{x}}_k^{\mathrm{DL,uc}}(n) = \mathbf{L}_k^H \mathbf{r}_k^{\mathrm{uc}}(n) \ . \tag{5.17}$$

5.3.3 Uplink data transmission

The final phase of the communication protocol consists of the uplink data transmission. In this phase the beamformer $\mathbf{L}_k$ is used as the precoder at the k-th MS, while the beamformer in Eq. (5.9) is used at the m-th AP as the postcoder combiner for the signal from the k-th MS. Since the MSs do not perform channel estimation, they just send their data symbols using the already defined trivial beamformer $\mathbf{L}_k$. The P_k-dimensional data vector to be transmitted by the k-th user in the n-th sample time is denoted by $\mathbf{x}_k^{\mathrm{UL}}(n)$. The signal received at the m-th AP in the n-th time sample is the N_{AP}-dimensional vector

$$\bar{\mathbf{y}}_m(n) = \sum_{k=1}^{K} \sqrt{\eta_k^{\mathrm{UL}}} \mathbf{G}_{k,m} \mathbf{L}_k \mathbf{x}_k^{\mathrm{UL}}(n) + \mathbf{w}_m(n) \ , \tag{5.18}$$

with η_k^{UL} is the uplink transmit power of the k-th MS.

- **CF massive MIMO architecture**

In the case of CF approach, all the APs participate to the decoding of the data sent by all the MSs. The m-th AP, thus, forms, for each $k = 1, \ldots, K$, the following statistics

$$\widetilde{\mathbf{y}}_{m,k}(n) = \left(\mathbf{L}_k^H \widehat{\mathbf{G}}_{k,m}^H \widehat{\mathbf{G}}_{k,m} \mathbf{L}_k\right)^{-1} \mathbf{L}_k^H \widehat{\mathbf{G}}_{k,m}^H \bar{\mathbf{y}}_m(n) = \widetilde{\mathbf{G}}_{k,m} \bar{\mathbf{y}}_m(n), \tag{5.19}$$

where $\widetilde{\mathbf{G}}_{k,m}$ is the following $P_k \times N_{\mathrm{AP}}$-dimensional matrix:

$$\widetilde{\mathbf{G}}_{k,m} = \left(\mathbf{L}_k^H \widehat{\mathbf{G}}_{k,m}^H \widehat{\mathbf{G}}_{k,m} \mathbf{L}_k\right)^{-1} \mathbf{L}_k^H \widehat{\mathbf{G}}_{k,m}^H . \tag{5.20}$$

The vectors $\widetilde{\mathbf{y}}_{m,k}(n)$, for all $k = 1, \ldots, K$, are then sent to the CPU via the backhaul link; the CPU, finally, forms the following soft estimates of the data vectors transmitted by the users:

$$\widehat{\mathbf{x}}_k^{\mathrm{UL,cf}}(n) = \sum_{m=1}^{M} \widetilde{\mathbf{y}}_{m,k}(n) , \quad k = 1, \ldots, K . \tag{5.21}$$

Note that only the soft estimates $\widetilde{\mathbf{y}}_{m,k}(n)$ are to be transmitted from the APs to the CPU, while channel estimates transmission is not required.

- **UC massive MIMO architecture**
  In this case, the signal transmitted by the k-th MS is decoded only by the APs in the set $\mathcal{M}(k)$. Otherwise stated, the m-th AP computes the statistics $\widetilde{\mathbf{y}}_{m,k}(n)$ only for the MSs in $\mathcal{K}(m)$. Accordingly, the CPU is able to perform the following soft estimates for the data sent by the K MSs in the system:

$$\widehat{\mathbf{x}}_k^{\mathrm{UL,uc}}(n) = \sum_{m \in \mathcal{M}(k)} \widetilde{\mathbf{y}}_{m,k}(n) , \quad k = 1, \ldots, K . \tag{5.22}$$

  Notice that in this case the backhaul overhed is reduced with respect to the CF case since each AP has to send only the soft estimates of the data received by its associated MSs.

5.4 Performance measures and downlink power control

Note that the CF approach can be obtained as a special case of UC by letting $N = K$, i.e., each AP serves all the K users in the system, so $\mathcal{M}(k) = \{1, \ldots, M\}$, $\forall k = 1, \ldots, K$. Following this approach, in the downlink power control here explained, the generic scalar coefficient ruling the power transmitted in downlink by the m-th AP for the k-th MS is denoted as $\eta_{k,m}^{\mathrm{DL}}$. From Eq. (5.16), the achievable rate in downlink for the user k can be written as

$$\mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}) = W \log_2 \left| \mathbf{I} + \mathbf{R}_k^{-1} \mathbf{A}_{k,k} \mathbf{A}_{k,k}^H \right| , \tag{5.23}$$

where

$$\mathbf{A}_{k,k} = \sum_{m \in \mathcal{M}(k)} \mathbf{L}_k^H \sqrt{\eta_{k,m}^{\mathrm{DL}}} \mathbf{G}_{k,m}^H \widehat{\mathbf{G}}_{k,m} \left(\widehat{\mathbf{G}}_{k,m}^H \widehat{\mathbf{G}}_{k,m} \right)^{-1} \mathbf{L}_k \,, \tag{5.24}$$

$$\mathbf{R}_k = \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathbf{A}_{k,j} \mathbf{A}_{k,j}^H \,, \tag{5.25}$$

$$\mathbf{A}_{k,j} = \mathbf{L}_k^H \sum_{m \in \mathcal{M}(j)} \sqrt{\eta_{j,m}^{\mathrm{DL}}} \mathbf{G}_{k,m}^H \widehat{\mathbf{G}}_{j,m} \left(\widehat{\mathbf{G}}_{j,m}^H \widehat{\mathbf{G}}_{j,m} \right)^{-1} \mathbf{L}_j \tag{5.26}$$

and $\boldsymbol{\eta}$ is the $KM \times 1$ vector collecting the transmit powers in downlink of all APs for all MSs. Notice that Eq. (5.23) represents the achievable rate in the case of perfect CSI, whereas it is just an approximation of the rate when channel estimates are considered. Nonetheless, it still represents a good indicator of the attainable performance, since the approximation error gets smaller as the reliability in the channel estimates increases. The rest of this section will be concerned with the optimization of the downlink transmit powers for the maximization of the system sum-rate and minimum users' rate, subject to maximum power constraints. Mathematically, the sum-rate maximization problem is formulated as the optimization program[15]:

$$\max_{\boldsymbol{\eta}} \sum_{k=1}^{K} \mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}) \tag{5.27a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \,, \forall\, m = 1, \ldots, M \tag{5.27b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall\, m = 1, \ldots, M,\ k = 1, \ldots, K, \tag{5.27c}$$

whereas the minimum-rate maximization problem is

$$\max_{\boldsymbol{\eta}} \min_{1 \leq k \leq K} \mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}) \tag{5.28a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \,, \forall\, m = 1, \ldots, M \tag{5.28b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall\, m = 1, \ldots, M,\ k = 1, \ldots, K. \tag{5.28c}$$

Both problems have non-concave objective functions, which makes their solution challenging. Moreover, even if the problems were concave, the large number of optimization variables, KM, would still pose a significant complexity challenge. In order to face these issues, the framework of successive lower-bound maximization is used, recently introduced in[16] [93], and briefly reviewed next.

[15] The procedure can be applied, with ordinary efforts, also in the case in which minimum-rate constraints are considered.

[16] In [93] the method is labeled successive upper-bound minimization, since minimization problems are considered.

5.4.1 Successive lower-bound maximization

The main idea of the method is to merge the tools of alternating optimization [94, Section 2.7] and sequential convex programming [95]. To elaborate, consider the generic optimization problem

$$\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \ , \tag{5.29}$$

with $f : \mathbb{R}^n \to \mathbb{R}$ a differentiable function, and $\mathcal{X}$ a compact set. As in the alternating optimization method, the successive lower-bound maximization partitions the variable space into M blocks, $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M)$, which are cyclically optimized one at a time, while keeping the other variable blocks fixed. This effectively decomposes Eq. (5.29) into M subproblems, with the generic subproblem stated as

$$\max_{\boldsymbol{x}_m} f(\boldsymbol{x}_m, \boldsymbol{x}_{-m}) \ , \tag{5.30}$$

with $\boldsymbol{x}_{-m}$ collecting all variable blocks except the m-th. It is proved in [94, Proposition 2.7.1] that iteratively solving Problem (5.30) monotonically improves the value of the objective of Problem (5.29), and converges to a first-order optimal point if the solution of Problem (5.30) is unique for any m, and if $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_M$, with $\boldsymbol{x}_m \in \mathcal{X}_m$ for all m. Clearly, alternating optimization proves useful when Problem (5.30) can be solved with minor complexity. If this is not the case, the successive lower-bound maximization method proposes to tackle Problem (5.30) by means of sequential convex programming. This does not guarantee to globally solve Problem (5.30), but can lead to a computationally feasible algorithm. Moreover, it is guaranteed to preserve the properties of the alternating optimization method [93]. The idea of sequential optimization is to tackle a difficult maximization problem by solving a sequence of easier maximization problems. To elaborate, let us denote by $g_i(\boldsymbol{x}_m)$ the i-th constraint of Problem (5.30), for $i = 1, \ldots, C$. Then, consider a sequence of approximate problems $\{\mathcal{P}_\ell\}_\ell$ with objectives $\{f_\ell\}_\ell$ and constraint functions $\{g_{i,\ell}\}_{i=1}^{C}$, such that the following three properties are fulfilled, for all ℓ:

(P1) $f_\ell(\boldsymbol{x}_m) \leq f(\boldsymbol{x}_m)$, $g_{i,\ell}(\boldsymbol{x}_m) \leq g_i(\boldsymbol{x}_m)$, for all i and $\boldsymbol{x}_m$;

(P2) $f_\ell(\boldsymbol{x}_m^{(\ell-1)}) = f(\boldsymbol{x}_m^{(\ell-1)})$, $g_{i,\ell}(\boldsymbol{x}_m^{(\ell-1)}) = g_i(\boldsymbol{x}_m^{(\ell-1)})$ with $\boldsymbol{x}_m^{(\ell-1)}$ the maximizer of $f_{\ell-1}$;

(P3) $\nabla f_\ell(\boldsymbol{x}_m^{(\ell-1)}) = \nabla f(\boldsymbol{x}_m^{(\ell-1)})$, $\nabla g_{i,\ell}(\boldsymbol{x}_m^{(\ell-1)}) = \nabla g_i(\boldsymbol{x}_m^{(\ell-1)})$.

In [95] (see also [93], [96]) it is shown that, subject to constraint qualifications, the sequence $\{f(\boldsymbol{x}_m^{(\ell)})\}_\ell$ of the solutions of the ℓ-th Problem $\mathcal{P}_\ell$, is monotonically increasing and converges. Moreover, every limit point of the sequence $\{\boldsymbol{x}_m^{(\ell)}\}_\ell$ attains a first-order optimal point of the original Problem (5.30). Thus, the sequential approach enjoys strong optimality properties, fulfilling at the same

time the monotonic improvement property for the objective function, and the Karush–Kuhn–Tucker (KKT) first-order optimality conditions for the original problem. Nevertheless, its applicability hinges on determining suitable lower bounds for the original objective to maximize, which fulfill all three properties **P1**, **P2**, **P3**, while at the same time leading to manageable optimization problems. In conclusion, the successive lower-bound maximization method can be formulated as variation of the alternating optimization method, in which each subproblem (5.30) is not globally solved, but instead is tackled by sequential optimization theory. It is proved in [93] that successive lower-bound maximization has the same optimality properties as the true alternating optimization method, under similar assumptions, even though each subproblem might not be globally solved[17].

5.4.2 Sum-rate maximization

Consider Problem (5.27) and define the variable blocks $\boldsymbol{\eta}_m$, $m = 1, \ldots, M$, collecting the transmit powers of AP m. Then, the sum-rate maximization with respect to the variable block $\boldsymbol{\eta}_m$ is cast as

$$\max_{\boldsymbol{\eta}_m} \sum_{k=1}^{K} \mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{-m}) \tag{5.31a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \tag{5.31b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall\, k \in \mathcal{K}_m \,. \tag{5.31c}$$

The complexity of Problem (5.31) is significantly lower than that of Problem (5.27), since only the M transmit powers of AP m are being optimized. Nevertheless, Problem (5.31) is still non-convex, which makes its solution difficult. Indeed, defining

$$\mathbf{A}_{k,j,m} = \mathbf{L}_k^H \mathbf{G}_{k,m}^H \widehat{\mathbf{G}}_{j,m} \left(\widehat{\mathbf{G}}_{j,m}^H \widehat{\mathbf{G}}_{j,m} \right)^{-1} \mathbf{L}_j \,, \tag{5.32}$$

the k-th user's achievable rate can be expressed as Eq. (5.33) at the top of next page, which can be seen to be non-concave, also with respect to only the variable block $\boldsymbol{\eta}_m$. Thus, following the successive lower-bound maximization, Problem (5.31) will be tackled by sequential optimization. To this end, it is necessary to derive a lower-bound of the objective of Problem (5.31), which fulfills Properties **P1**, **P2**, and **P3**, while at the same time leading to a simple optimization problem. To this end, the following lemma proves useful.

Lemma 1. *The function $f : (x, y) \in \mathbb{R}^2 \to \sqrt{xy}$ is jointly concave in x and y, for $x, y > 0$.*

[17] Of course, this holds provided the additional assumption of the sequential method are fulfilled in each iteration

$$\mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}) = \underbrace{W \log_2 \left| \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k + \sum_{j=1}^{K} \sum_{m} \sum_{\ell} \sqrt{\eta_{j,m}^{\mathrm{DL}} \eta_{j,\ell}^{\mathrm{DL}}} \mathbf{A}_{k,j,m} \mathbf{A}_{k,j,\ell}^H \right|}_{g_1(\boldsymbol{\eta})} \tag{5.33}$$

$$- \underbrace{W \log_2 \left| \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k + \sum_{\substack{j=1 \\ j \neq k}}^{K} \sum_{m} \sum_{\ell} \sqrt{\eta_{j,m}^{\mathrm{DL}} \eta_{j,\ell}^{\mathrm{DL}}} \mathbf{A}_{k,j,m} \mathbf{A}_{k,j,\ell}^H \right|}_{g_2(\boldsymbol{\eta})} .$$

Proof: The proof follows upon computing the Hessian of $\sqrt{xy}$ and noting that it is negative semi-definite.

Lemma 1, coupled with the facts that the function $\log_2 |(\cdot)|$ is matrix-increasing, and that summation preserves concavity, implies that the rate function in Eq. (5.33) is the difference of two concave functions. This observation is instrumental for the derivation of the desired lower-bound. Indeed, recalling that any concave function is upper-bounded by its Taylor expansion around any given point $\boldsymbol{\eta}_{m,0}$, a concave lower-bound of $\mathcal{R}_k$ is obtained as

$$\begin{aligned} \mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}) &= g_1(\boldsymbol{\eta}_m) - g_2(\boldsymbol{\eta}_m) \\ &\geq g_1(\boldsymbol{\eta}_m) - g_2(\boldsymbol{\eta}_{m,0}) - \nabla_{\boldsymbol{\eta}_m}^T g_2|_{\boldsymbol{\eta}_{m,0}} (\boldsymbol{\eta}_m - \boldsymbol{\eta}_{m,0}) \\ &= \widetilde{\mathcal{R}}_k^{\mathrm{DL}}(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{m,0}) \, . \end{aligned} \tag{5.34}$$

Moreover, it is easy to check that $\widetilde{\mathcal{R}}_k$ fulfills by construction also properties **P2** and **P3** with respect to $\mathcal{R}_k$. Thus, Problem (5.31) can be tackled by the sequential optimization framework, upon defining the ℓ-th problem of the sequence, $\mathcal{P}_\ell$, as the convex optimization program:

$$\max_{\boldsymbol{\eta}_m} \sum_{k=1}^{K} \widetilde{\mathcal{R}}_k(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{m,0}, \boldsymbol{\eta}_{-m}) \tag{5.35a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}_m} \eta_{k,m} \leq P_{\max,m} \tag{5.35b}$$

$$\eta_{k,m} \geq 0 \, , \forall \, k \in \mathcal{K}_m \tag{5.35c}$$

For any $\boldsymbol{\eta}_{m,0}$, Problem (5.35) can be solved by means of standard convex optimization theory, since the objective is concave, and the constraints are affine. Additionally, with this formulation, a reduced variables set is to be optimized, with respect to the original Problem (5.27). The resulting power control procedure can be stated as in Algorithm 4. Moreover, based on the general theory reviewed in Section 5.4.1, the following result holds.

Proposition 1. *After each iteration in Line 6 of Algorithm 4, the sum-rate value* $\sum_{k=1}^{K} \mathcal{R}_k^{\mathrm{DL}}$

Algorithm 4 Sum-rate maximization

Set $i = 0$ and choose any feasible $\boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_M$;
repeat
  for $m = 1 \rightarrow M$ **do**
    repeat
      Choose any feasible $\boldsymbol{\eta}_{m,0}$;
      Let $\boldsymbol{\eta}_m^*$ be the solution of Problem (5.35);
      $\boldsymbol{\eta}_{m,0} = \boldsymbol{\eta}_m^*$;
    until convergence
    $\boldsymbol{\eta}_m = \boldsymbol{\eta}_m^*$;
  end for
until convergence

is not decreased, and the resulting sequence $\left\{ \sum_{k=1}^{K} \mathcal{R}_k^{\mathrm{DL}} \right\}$ *converges. Moreover, every limit point of the sequence* $\{\boldsymbol{\eta}_m\}_m$ *fulfills the KKT first-order optimality conditions of Problem* (5.31).

Two remarks are now in order. First of all an extreme case of Algorithm 4 provides that only one variable block is used, namely optimizing all of the transmit powers simultaneously. In this scenario, Algorithm 4 reduces to a pure instance of sequential optimization, and no alternating optimization is required. Nevertheless, as already mentioned, the complexity of this approach seems prohibitive for large M and K. Then, another extreme case is that in which the KM transmit powers $\eta_{k,m}$ are optimized one at a time, thus leading to considering KM variable blocks. The advantage of this approach is that each subproblem (5.35) would have only one optimization variable, and thus could be solved in semi-closed form. This brings drastic computational complexity savings and proves to be useful especially in the CF scenario, since in this case each variable block $\boldsymbol{\eta}_m$ always has dimension K.

Computational complexity: The complexity of Algorithm 4 depends on the complexity of Problem (5.35), and on how many such problems must be solved before convergence. As for Problem (5.35), it is a convex problem and as such its complexity is polynomial in the number of variables, even though the specific degree of the polynomial is not known. The best available upper-bound for generic convex problems is provided in [97], and states that the complexity of any convex problem scales at most with the fourth power of the number of variables. Instead, as for the number of iterations required for the "while" loops in Algorithm 4 to reach convergence, no closed-form result is currently available. Nevertheless, defining by I_O and I_S the number of iterations for the outer and inner "while" loops to converge, the overall complexity of Algorithm 4 can be upper-bounded by $\mathcal{O}\left(I_O I_S \sum_{m=1}^{M} |\mathcal{K}_m|^4\right)$.

5.4.3 Minimum-rate maximization

Consider Problem (5.28). Following similar steps as in Section 5.4.2, Problem (5.28) with respect to variable block $\boldsymbol{\eta}_m$ becomes

$$\max_{\boldsymbol{\eta}_m} \min_{1\leq k\leq K} \mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{-m}) \tag{5.36a}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \tag{5.36b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall k \in \mathcal{K}_m \tag{5.36c}$$

Besides the difficulties already encountered in the sum-rate scenario, Problem (5.36) poses the additional challenge of having a non-differentiable objective due to the $\min(\cdot)$ operator. To circumvent this issue, Problem (5.36) can be equivalently reformulated as the program:

$$\max_{\boldsymbol{\eta}_m, t} \; t \tag{5.37a}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \tag{5.37b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall\, k \in \mathcal{K}_m \tag{5.37c}$$

$$\mathcal{R}_k^{\mathrm{DL}}(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{-m}) \geq t \,, \forall\, k = 1, \ldots, K \,. \tag{5.37d}$$

At this point, it is possible to tackle Problem (5.37) by the sequential method. Leveraging again the bound in Eq. (5.34) leads to considering the approximate problem

$$\max_{\boldsymbol{\eta}_m, t} \; t \tag{5.38a}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}_m} \eta_{k,m}^{\mathrm{DL}} \leq P_{\max,m} \tag{5.38b}$$

$$\eta_{k,m}^{\mathrm{DL}} \geq 0 \,, \forall\, k \in \mathcal{K}_m \tag{5.38c}$$

$$\widetilde{\mathcal{R}}_k^{\mathrm{DL}}(\boldsymbol{\eta}_m, \boldsymbol{\eta}_{m,0}, \boldsymbol{\eta}_{-m}) \geq t \,, \forall\, k = 1, \ldots, K \,. \tag{5.38d}$$

For any $\boldsymbol{\eta}_{m,0}$, Problem (5.38) can be solved by means of standard convex optimization theory, since the objective is linear, and the constraints are all convex. The resulting power control procedure can be stated as in Algorithm 5, which enjoys similar properties as Algorithm 4.

Computational complexity: Following similar arguments as done in Section 5.4.2 for the sum-rate case, the complexity of Algorithm 5 is upper-bounded by $\mathcal{O}\left(I_O I_S \sum_{m=1}^{M} (|\mathcal{K}_m|+1)^4\right)$, where it has been accounted for the fact that the number of variables in the generic Problem (5.38) is $|\mathcal{K}_m|+1$ due to the presence of the auxiliary variable t.

Algorithm 5 Minimum-rate maximization

Set $i = 0$ and choose any feasible $\boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_M$;
repeat
  for $m = 1 \rightarrow M$ **do**
    repeat
      Choose any feasible $\boldsymbol{\eta}_{m,0}$;
      Let $\boldsymbol{\eta}_m^*$ be the solution of Problem (5.38);
      $\boldsymbol{\eta}_{m,0} = \boldsymbol{\eta}_m^*$;
    until convergence
    $\boldsymbol{\eta}_m = \boldsymbol{\eta}_m^*$;
  end for
until convergence

5.5 Uplink power control

Considering again the CF approach as a special case of the UC one, from Eq. (5.22), for the uplink, upon defining

$$\mathbf{B}_{k,j} = \sum_{m \in \mathcal{M}(k)} \widetilde{\mathbf{G}}_{k,m} \mathbf{G}_{j,m} \mathbf{L}_j \tag{5.39}$$

it is obtained that the approximation (for the case of non-perfect channel knowledge) of the achievable rate for the k-th user is

$$\mathcal{R}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}}) = \log_2 \left| \mathbf{I}_{P_k} + \eta_k^{\mathrm{UL}} \widetilde{\mathbf{R}}_k^{-1} \mathbf{B}_{k,k} \mathbf{B}_{k,k}^H \right| , \tag{5.40}$$

where

$$\widetilde{\mathbf{R}}_k = \sum_{\substack{j=1 \\ j \neq k}}^{K} \eta_j^{\mathrm{UL}} \mathbf{B}_{k,j} \mathbf{B}_{k,j}^H + \sigma_w^2 \sum_{m \in \mathcal{M}(k)} \widetilde{\mathbf{G}}_{k,m} \widetilde{\mathbf{G}}_{k,m}^H , \tag{5.41}$$

and $\widetilde{\boldsymbol{\eta}}$ is the $K \times 1$ vector collecting the transmit uplink powers of all MSs. Then, the sum-rate maximization problem is stated as

$$\max_{\widetilde{\boldsymbol{\eta}}} \sum_{k=1}^{K} \mathcal{R}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}}) \tag{5.42a}$$

$$\text{s.t. } 0 \leq \eta_k^{\mathrm{UL}} \leq P_{\max,k} \, \forall \, k = 1, \ldots, K , \tag{5.42b}$$

while the minimum-rate maximization problem is stated as

$$\max_{\widetilde{\boldsymbol{\eta}}} \min_{1,\ldots,K} \mathcal{R}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}}) \tag{5.43a}$$

$$\text{s.t. } 0 \leq \eta_k^{\mathrm{UL}} \leq P_{\max,k} \, \forall \, k = 1, \ldots, K . \tag{5.43b}$$

Based on Eq. (5.40), it is possible to develop power control algorithms for sum-rate and minimum-rate maximization, leveraging the sequential optimization framework, as done in the downlink scenario. Indeed, also in the uplink case, it

is observed that the k-th user's rate in Eq. (5.40) can be written as the difference of two concave functions, namely

$$\mathcal{R}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}}) = \underbrace{W\log_2\left|\mathbf{G}_k + \sum_{j=1}^{K}\eta_j^{\mathrm{UL}}\mathbf{B}_{k,j}\mathbf{B}_{k,j}^H\right|}_{\widetilde{g}_1(\widetilde{\boldsymbol{\eta}})} - \underbrace{W\log_2\left|\mathbf{G}_k + \sum_{\substack{j=1\\ j\neq k}}^{K}\eta_j^{\mathrm{UL}}\mathbf{B}_{k,j}\mathbf{B}_{k,j}^H\right|}_{\widetilde{g}_2(\widetilde{\boldsymbol{\eta}})}, \quad (5.44)$$

wherein $\mathbf{G}_k = \sigma_w^2 \sum_{m\in\mathcal{M}(k)} \widetilde{\mathbf{G}}_{k,m}\widetilde{\mathbf{G}}_{k,m}^H$ for all $k = 1,\ldots,K$. Now, it is clear that both $\widetilde{g}_1(\cdot)$ and $\widetilde{g}_2(\cdot)$ are concave functions of $\widetilde{\boldsymbol{\eta}}$ and thus Eq. (5.44) shows that the k-th user's rate can be once again written as the difference of two concave functions. As a consequence, for all $k = 1,\ldots,K$, a lower-bound of the k-th user's rate, which fulfills all three properties of the sequential optimization method, say $\widetilde{\mathcal{R}}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}})$, is given by Eq. (5.34), in which $\widetilde{g}_1$ and $\widetilde{g}_2$ take the expression in Eq. (5.44) above.

Remark 1. In the downlink case the number of optimization variables was KM, with $M > K$, and this made it convenient, for complexity reasons, to partition the variable space into multiple blocks of variables that were alternatively optimized. On the other hand, in the uplink case there are only K variables, and this makes it practically feasible to consider only one variable block, thus optimizing all variables at the same time. In the sequel, the focus will be on this case, but we stress that, if desired, the optimization algorithms can be straightforwardly extended to the scenario in which multiple optimization blocks are defined and iteratively optimized.

Keeping Remark 1 in mind, both sum-rate maximization and minimum-rate maximization can be performed by similar algorithms as Algorithms 4 and 5, respectively, in which the auxiliary problem to be solved within each iteration are stated as

$$\max_{\widetilde{\boldsymbol{\eta}}} \sum_{k=1}^{K}\widetilde{\mathcal{R}}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}},\widetilde{\boldsymbol{\eta}}_0) \quad (5.45a)$$

$$\text{s.t. } 0 \le \eta_k^{\mathrm{UL}} \le P_{\max,k}\ \forall\, k = 1,\ldots,K \quad (5.45b)$$

for sum-rate maximization, and as

$$\max_{\widetilde{\boldsymbol{\eta}}_m,t}\ t \quad (5.46a)$$

$$\text{s.t. } 0 \le \eta_k^{\mathrm{UL}} \le P_{\max,k}\ \forall\, k = 1,\ldots,K \quad (5.46b)$$

$$\widetilde{\mathcal{R}}_k^{\mathrm{UL}}(\widetilde{\boldsymbol{\eta}},\widetilde{\boldsymbol{\eta}}_0) \ge t\ ,\forall\, k = 1,\ldots,K\ , \quad (5.46c)$$

for minimum-rate maximization. Similar optimality properties as in the downlink case hold.

Computational complexity: Following similar arguments as for the downlink scenario, the computational complexity of the proposed approach in the uplink scenario is upper-bounded by $\mathcal{O}\left(I_S K^4\right)$ in the sum-rate maximization case, and $\mathcal{O}\left(I_S (K+1)^4\right)$ in the minimum-rate maximization case, where it has been accounted for the fact that, as stated in Remark 1, only one variable block is considered in the uplink scenario, thus removing the outer loop that is instead present in Algorithms 4 and 5 in the downlink scenario.

5.6 Advanced signal processing techniques for UC architecture

In this Section advanced signal processing techniques are proposed in order to improve the performance of the UC architecture. These techniques are the linear MMSE channel estimation, the local PZF pre-coding on the downlink and SIC on the uplink. The proposed schemes can be implemented locally, i.e., the channel estimates and the beamformers at the APs are computed and exploited locally, with no need to exchange information with a CPU.

5.6.1 Linear MMSE channel estimation

To perform linear MMSE channel estimation, the $(N_{\mathrm{AP}}\tau_p)$-dimensional vector $\check{\mathbf{y}}_m = \mathrm{vec}\left(\mathbf{Y}_m\right)$ is considered, with the matrix $\mathbf{Y}_m$ as in Eq. (5.5), which can be shown to be expressed as

$$\check{\mathbf{y}}_m = \sum_{k=1}^{K} \sqrt{p_k} \left(\mathbf{\Phi}_k^T \otimes \mathbf{I}_{N_{\mathrm{AP}}}\right) \check{\mathbf{g}}_{k,m} + \check{\mathbf{w}}_m \ , \tag{5.47}$$

where $\check{\mathbf{g}}_{k,m} = \mathrm{vec}\left(\mathbf{G}_{k,m}\right)$ is an $(N_{\mathrm{AP}} N_{\mathrm{MS}})$-dimensional vector and $\check{\mathbf{w}}_m = \mathrm{vec}\left(\mathbf{W}_m\right)$ is an $(N_{\mathrm{AP}}\tau_p)$-dimensional vector. Neglecting the correlation of the shadow fading and using the simplifying assumption $\mathbb{E}\left[\check{\mathbf{g}}_{k,m}\check{\mathbf{g}}_{k,m}^H\right] = \mathbf{I}_{N_{\mathrm{AP}}N_{\mathrm{MS}}}$, the linear MMSE estimate of $\check{\mathbf{g}}_{k,m}$ is obtained as [22]:

$$\widehat{\check{\mathbf{g}}}_{k,m} = \mathbf{R}_{k,m}\mathbf{R}_m^{-1}\check{\mathbf{y}}_m \ , \tag{5.48}$$

where $\mathbf{R}_{k,m} = \sqrt{p_k}\mathbf{\Phi}_k^* \otimes \mathbf{I}_{N_{\mathrm{AP}}}$, and

$$\mathbf{R}_m = \sum_{j=1}^{K} p_j \left(\mathbf{\Phi}_j^T \mathbf{\Phi}_j^* \otimes \mathbf{I}_{N_{\mathrm{AP}}}\right) + \sigma_w^2 \mathbf{I}_{N_{\mathrm{AP}}\tau_p} \ . \tag{5.49}$$

5.6.2 PZF downlink pre-coding

The multiplexing order for all the communications is assumed equal for all the users in the system, i.e., $P_k = P, \ \forall k = 1, \ldots, K$ and, consequently the channel independent beamformer are equal for all the users, i.e., $\mathbf{L}_k = \mathbf{L}, \ \forall k = 1, \ldots, K$.

The signal transmitted by the m-th AP in the n-th interval can be expressed as Eq. (5.15). The generic k-th MS receives signal contributions from all the APs and the observable vector is expressed as Eq. (5.16). Based on the observation of the vector $\mathbf{r}_k^{\mathrm{uc}}(n)$, a soft estimate of the data symbols $\mathbf{x}_k^{\mathrm{DL}}(n)$ is obtained at the k-th MS as

$$\widehat{\mathbf{x}}_k^{\mathrm{DL}}(n) = \mathbf{L}^H \mathbf{r}_k^{\mathrm{uc}}(n) \ . \tag{5.50}$$

In order to design $\mathbf{Q}_{k,m}$, we assume that each AP performs a local PZF downlink pre-coding, so that it does not interfere with a certain number, say S, of MSs[18]. Another objective of this beamforming scheme is to ensure that the MSs will be able to receive data with no information on the channel state. We denote by $\mathcal{S}(k,m)$ the set, of cardinality S, containing the MSs towards which to null the interference in the design of the precoder $\mathbf{Q}_{k,m}^{\mathrm{DL}}$. The N_{AP}-dimensional p-th column of the matrix $\mathbf{Q}_{k,m}^{\mathrm{DL}}$ can be thus chosen in order to fulfill the following conditions:

$$\begin{cases} \left(\mathbf{L}^H \widehat{\mathbf{G}}_{k,m}^H\right)_{(p,:)} \mathbf{Q}_{k,m}(:,p) = 1 \ , \\ \left(\mathbf{L}^H \widehat{\mathbf{G}}_{k,m}^H\right)_{(\bar{p},:)} \mathbf{Q}_{k,m}(:,p) = 0 \ , \\ \left(\mathbf{L}^H \widehat{\mathbf{G}}_{j,m}^H\right)_{(\ell,:)} \mathbf{Q}_{k,m}(:,p) = 0 \ , \\ \forall \ell = 1,\ldots,P \ , \forall j \in \mathcal{S}(k,m) \ , \end{cases} \tag{5.51}$$

where the generic $(\mathbf{A})_{(\bar{p},:)}$ denotes a matrix contains all the rows of $\mathbf{A}$ except the p-th. The above conditions can be easily fulfilled provided that $N_{\mathrm{AP}} \geq SP + P - 1$. Regarding the set $\mathcal{S}(k,m)$, it assumed that it contains the S MSs, served by the m-th AP, with the higher Frobenius norm, except the k-th.

5.6.3 SIC procedure for the uplink

Regarding the design of the SIC procedure on the uplink, starting from the set $\mathcal{K}(m)$ each AP uses the following sorting operator:

$$O_m : \mathcal{K}(m) \rightarrow \{1,\ldots,K\} \ , \tag{5.52}$$

where $O_m(k)$ is the index of the MS that is in the k-th position when MSs are ordered in terms of decreasing Frobenius norm of the channel estimates available at the m-th AP. Otherwise stated, the following relation holds:

$$\|\widehat{\mathbf{G}}_{O_m(1),m}\|_F \geq \|\widehat{\mathbf{G}}_{O_m(2),m}\|_F \geq \ldots \geq \|\widehat{\mathbf{G}}_{O_m(N),m}\|_F \tag{5.53}$$

The signal received at the m-th AP in the n-th time sample is $\bar{\mathbf{y}}_m(n)$ in Eq. (5.18). In order to illustrate the SIC processing at each AP, we start considering the processing for the MS $O_m(1)$, that, according to the definition in Eq. (5.52), is the MS with the largest channel Frobenius norm at the m-th AP, and is thus the MS

[18]Indeed it is not possible to null the interference towards all the MSs due to the fact that the AP may not be equipped with the needed number of antennas to perform this operation.

whose data are detected first. Upon defining the $P \times N_{\mathrm{AP}}$-dimensional matrix $\widetilde{\mathbf{G}}_{k,m}$ in Eq. (5.20), the m-th AP forms the following statistic for the MS $O_m(1)$:

$$\widetilde{\mathbf{u}}_{O_m(1),m}(n) = \widetilde{\mathbf{G}}_{O_m(1),m}\bar{\mathbf{y}}_m(n)\,. \tag{5.54}$$

Then, when forming the sufficient statistic for the data transmitted by the MS $O_m(2)$, the m-th AP cancels the contribution of the estimate of the signal transmitted from the MS $O_m(1)$ as follows:

$$\widetilde{\mathbf{u}}_{O_m(2),m}(n) = \widetilde{\mathbf{G}}_{O_m(2),m}\left(\bar{\mathbf{y}}_m(n) - \widehat{\mathbf{G}}_{O_m(1),m}\mathbf{L}\widetilde{\mathbf{u}}_{O_m(1),m}(n)\right). \tag{5.55}$$

In general, at the ℓ-th stage, the m-th AP cancels the contribution of all the estimates of the signals transmitted from the MSs $O_m(1),\ldots,O_m(\ell-1)$, so it performs

$$\widetilde{\mathbf{u}}_{O_m(\ell),m}(n) = \widetilde{\mathbf{G}}_{O_m(\ell),m}\left(\bar{\mathbf{y}}_m(n) - \sum_{j=1}^{\ell-1}\widehat{\mathbf{G}}_{O_m(j),m}\mathbf{L}\widetilde{\mathbf{u}}_{O_m(j),m}(n)\right). \tag{5.56}$$

Through simple algebra, it can be shown that Eq. (5.56) can be expressed as

$$\widetilde{\mathbf{u}}_{O_m(\ell),m}(n) = \widetilde{\mathbf{G}}_{O_m(\ell),m}\left(\prod_{j=1}^{\ell-1}\mathbf{I}_{N_{\mathrm{AP}}} - \widehat{\mathbf{G}}_{O_m(j),m}\mathbf{L}\widetilde{\mathbf{G}}_{O_m(j),m}\right)\bar{\mathbf{y}}_m(n). \tag{5.57}$$

The vectors containing the estimates of the symbols transmitted by each MSs, after the SIC implemented at each AP, are then sent to the CPU via the backhaul link; the CPU, finally, forms the following soft estimates of the data vectors transmitted by the users:

$$\widehat{\mathbf{x}}_k^{\mathrm{UL,SIC}}(n) = \sum_{m\in\mathcal{M}(k)}\widetilde{\mathbf{u}}_{m,k}(n)\,, \quad k=1,\ldots,K\,. \tag{5.58}$$

Notice that in this case the backhaul overhead is reduced with respect to the full CF case since each AP has to send only the soft estimates of the data received by its associated MSs. Another sorting operator is now defined, one for each user, i.e.,

$$\widetilde{O}_k : \{1,\ldots,M\} \to \{0,1,\ldots,N\}\,, \tag{5.59}$$

where $\widetilde{O}_k(m)$ denotes the stage in which the m-th AP serves the MS k-th; the stage "0" means that the AP m-th does not decode the k-th MS. It can be shown that, using Eq. (5.57) and the sorting operator in Eq. (5.59), Eq. (5.22) can be expressed as

$$\widehat{\mathbf{x}}_k^{\mathrm{UL,SIC}}(n) = \sum_{m\in\mathcal{M}(k)}\mathbf{C}_{k,m}\bar{\mathbf{y}}_m(n), \tag{5.60}$$

with $k = 1, \ldots, K$, where $\mathbf{C}_{k,m}$ is a $P \times N_{\mathrm{AP}}$-dimensional matrix defined as

$$\mathbf{C}_{k,m} = \widetilde{\mathbf{G}}_{k,m} \prod_{j=1}^{\widetilde{O}_k(m)-1} \left[\mathbf{I}_{N_{\mathrm{AP}}} - \widehat{\mathbf{G}}_{O_m(j),m} \mathbf{L} \widetilde{\mathbf{G}}_{O_m(j),m} \right]. \tag{5.61}$$

Accordingly, the soft estimate of the data-symbol transmitted from the k-th MS in the n-th time epoch can be finally written as

$$\begin{aligned} \widehat{\mathbf{x}}_k^{\mathrm{UL,SIC}}(n) &= \sqrt{\eta_k^{\mathrm{UL}}} \sum_{m \in \mathcal{M}(k)} \mathbf{C}_{k,m} \mathbf{G}_{k,m} \mathbf{L} \mathbf{x}_k^{\mathrm{UL}}(n) \\ &+ \sum_{\substack{\ell=1 \\ \ell \neq k}}^{K} \sqrt{\eta_\ell^{\mathrm{UL}}} \sum_{m \in \mathcal{M}(k)} \mathbf{C}_{k,m} \mathbf{G}_{\ell,m} \mathbf{L} \mathbf{x}_\ell^{\mathrm{UL}}(n) + \sum_{m \in \mathcal{M}(k)} \mathbf{C}_{k,m} \mathbf{w}_m(n). \end{aligned} \tag{5.62}$$

Based on Eq. (5.62), the k-th user uplink achievable rate is expressed as

$$\mathcal{R}_k^{\mathrm{UL,SIC}} = W \log_2 \left| \mathbf{I}_P + \sqrt{\eta_k^{\mathrm{UL}}} \check{\mathbf{R}}_k^{-1} \widetilde{\mathbf{B}}_{k,k} \widetilde{\mathbf{B}}_{k,k}^H \right| , \tag{5.63}$$

with

$$\check{\mathbf{R}}_k = \sum_{\substack{j=1 \\ j \neq k}}^{K} \sqrt{\eta_j^{\mathrm{UL}}} \widetilde{\mathbf{B}}_{k,j} \widetilde{\mathbf{B}}_{k,j}^H + \sigma_z^2 \sum_{m \in \mathcal{M}(k)} \mathbf{C}_{k,m} \mathbf{C}_{k,m}^H , \tag{5.64}$$

the covariance matrix of the interfering terms, and

$$\widetilde{\mathbf{B}}_{k,j} = \sum_{m \in \mathcal{M}(k)} \mathbf{C}_{k,m} \mathbf{G}_{j,m} \mathbf{L} . \tag{5.65}$$

Similarly to Eqs. (5.23) and (5.40), Eq. (5.63) represents the achievable rate in the case of perfect CSI, whereas it is just an approximation of the rate for the case in which channel estimates are considered. Nonetheless, it still represents a good indicator of the attainable performance, since the approximation error gets smaller as the reliability in the channel estimates increases.

5.7 Numerical results

In the simulation setup, a communication bandwidth of $W = 20$ MHz centered over the carrier frequency $f_0 = 1.9$ GHz is considered. The antenna height at the AP is 15 m and at the MS is 1.65 m. The standard deviation of the shadow fading is $\sigma_{\mathrm{sh}} = 8$ dB, the parameters for the three slope path loss model in Eq. (3.3) are $d_1 = 50$ m and $d_0 = 10$ m, the parameter δ in Eq. (5.2) is 0.5 and the correlation distance in Eq. (5.3) is $d_{\mathrm{decorr}} = 100$ m. The additive thermal noise is assumed to have a power spectral density of -174 dBm/Hz, while the front-end receiver at the AP and at the MS is assumed to have a noise figure of 9 dB. In order to emulate an infinite area and to avoid boundary effects, the square area

is wrapped around. The considered setup is taken from the originally formulated version of CF massive MIMO presented in [35]. The shown results come from an average over 100 random scenario realizations with independent MSs and APs locations and channels. It is quantitatively studied and compared the performance of the CF and UC massive MIMO architectures. A square area of 1000×1000 (square meters) as in [35] is considered; this area may be representative of a typical crowded environment where ultra-dense deployment of APs will be needed. The numbers of antennas at the APs and at the MSs are assumed $N_{\mathrm{AP}} = 4$ and $N_{\mathrm{MS}} = 2$, respectively, and the multiplexing order per user is $P_k = 2,\ \forall k = 1, \ldots, K$. For the PM channel estimation, it is used maximum-length-sequences (pseudo-noise) with length τ_p and the uplink transmit power for channel estimation is $p_k = 50$ mW for each antenna, $\forall k = 1, \ldots, K$. The performance measures here considered are the outage probability in the UC approach, i.e the probability that a MSs is not served by any APs in the system, the achievable rate per user, and the sum-rate of the system in uplink and downlink, both measured in bit/s. The considered power allocation rules will be compared for benchmarking with the uniform power allocation strategy. For the uniform power allocation in downlink, it is simply assumed that each AP uniformly divides its maximum power among the users that it serves in the system. So, for the CF massive MIMO architecture is

$$\eta_{k,m}^{\mathrm{DL,cf}} = \frac{P_{\mathrm{max},m}}{K \mathrm{tr}\left(\mathbf{Q}_{k,m}\mathbf{Q}_{k,m}^{H}\right)}\ , \tag{5.66}$$

and for the UC massive MIMO architecture is

$$\eta_{k,m}^{\mathrm{DL,uc}} = \begin{cases} \dfrac{P_{\mathrm{max},m}}{\mathrm{card}\left[\mathcal{K}(m)\right] \mathrm{tr}\left(\mathbf{Q}_{k,m}\mathbf{Q}_{k,m}^{H}\right)} & \text{if } \ k \in \mathcal{K}(m) \\ 0 & \text{otherwise .} \end{cases} \tag{5.67}$$

For the uniform power allocation in uplink, each MSs transmits with its maximum power, so in the CF and UC massive MIMO architecture

$$\eta_{k}^{\mathrm{UL,cf}} = \eta_{k}^{\mathrm{UL,uc}} = \frac{P_{\mathrm{max},k}}{N_{\mathrm{MS}}}\ ,\ \forall k = 1, \ldots, K. \tag{5.68}$$

In the following results, the maximum power available at each AP is 200 mW, i.e., $P_{\mathrm{max},m} = 200$ mW, $\forall\, m$ and the maximum power available at each MS is 100 mW, i.e., $P_{\mathrm{max},k} = 100$ mW, $\forall k = 1, \ldots, K$.

TABLE 5.1: Probability of observing an unserved MS in the UC approach for $M = 80$, $K = 15$ and $\tau_p = 16$

N	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Perfect CSI	0.12	0.0274	0.0083	0.0031	0.0013	$5.57\,10^{-4}$	$2.86\,10^{-4}$
Partial CSI	0.0283	0.0010	$8\,10^{-5}$	$6.66\,10^{-6}$	0	0	0

N	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Perfect CSI	$1.2\,10^{-4}$	$6.66\,10^{-5}$	$3.33\,10^{-5}$	0	0	0	0	0
Partial CSI	0	0	0	0	0	0	0	0

TABLE 5.2: Probability of observing an unserved MS in the UC approach for $M = 50$, $K = 5$ and $\tau_p = 8$

N	**1**	**2**	**3**	**4**	**5**
Perfect CSI	0.048	0.0048	$5.6\,10^{-4}$	10^{-4}	0
Partial CSI	0.0015	$4\,10^{-5}$	0	0	0

The first performance measure considered is the outage probability in the UC approach. Indeed, one of the possible drawbacks in the considered AP-MS association rule is that it may happen that a MS does not end up associated with any AP. After that the association AP-MS has been made in the UC approach, possibly unserved MS might be associated to the closest AP to solve thus this problem. On the other hand, it is showed that the probability to have unserved MSs is very low, so having an unserved MS is a rare event. In Tables 5.1 and 5.2 the probability (estimated over 10^7 realizations) of having an unserved MS versus N is reported, for a high density and a low density scenario. In the high density scenario it is assumed $M = 80$, $K = 15$, and $\tau_p = 16$, while in the low density scenario $M = 50$, $K = 5$, and $\tau_p = 8$ is considered. Note that increasing the value of N, in both the cases of perfect and partial (i.e., estimated) CSI, the outage probability decreases and reaches the value 0, i.e., all the users are served at least by one AP in the system. It can be noted that in the case of partial CSI the outage probability reaches the value 0 faster than in the case of perfect CSI; this can be due to the fact that the randomness introduced by the noise and the interference in the channel estimated reduces the probability of having an unserved MS.

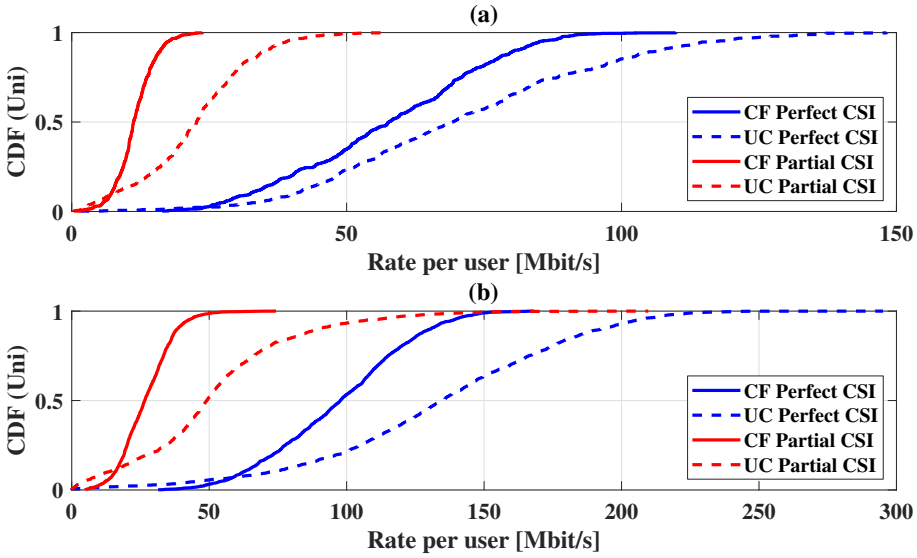

FIGURE 5.2: CDF of rate per user in downlink with uniform power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

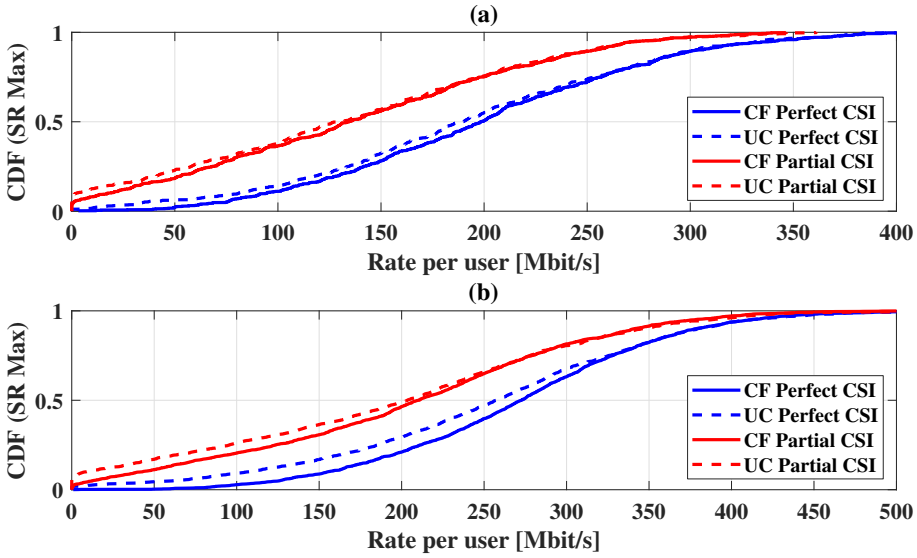

FIGURE 5.3: CDF of rate per user in downlink with sum-rate maximizing power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

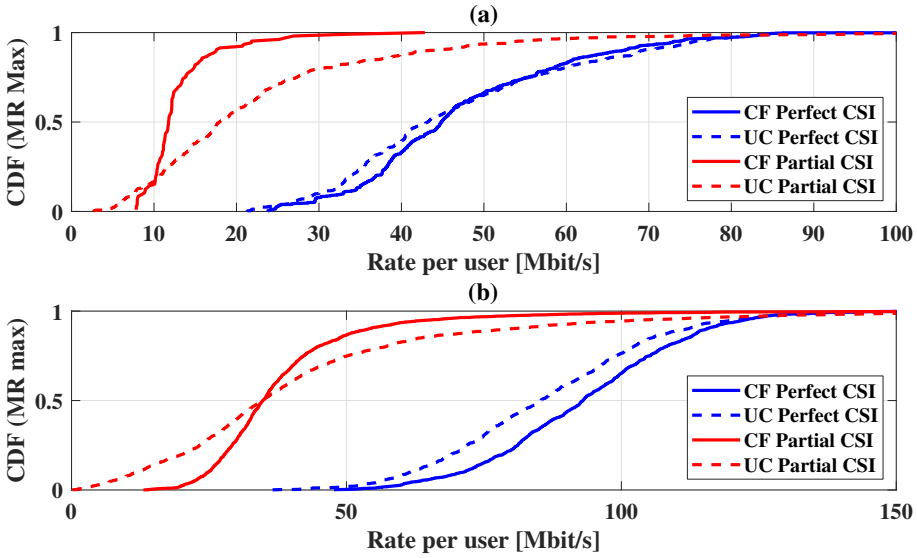

FIGURE 5.4: CDF of rate per user in downlink with minimum-rate maximizing power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

Then the performance measures considered are the achievable rates. Fig. 5.2 shows the cumulative distribution functions (CDFs) of the rate per user in downlink for the CF and UC approaches for the case in which uniform power allocation (Uni) is used. Fig. 5.3 reports the CDFs of the rate per user in downlink for the CF and UC approaches for the case in which sum-rate maximizing power allocation (SR Max) is used and Fig. 5.4 reports CDFs of the rate per user in downlink for the CF and UC approaches for the case in which minimum-rate maximizing power allocation (MR Max) is used. Both the cases of perfect CSI and partial CSI are considered. In Figs. 5.2-5.4 the results of the rate per user in the high and low density scenarios are compared, for the high density scenario $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$, while for the low density scenario $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$. Inspecting these figures, the following observation can be pointed out:

- In the low-rate region of the CDFs, the CF approach generally outperforms the UC approach; this part of the CDF reports the performance of the unlucky MSs that have very bad channels towards all the APs. These MSs take advantage of the CF deployment since they are served by a larger number of APs (compared to the UC case) and this explains the superiority of the CF approach for these MSs.

- The curves corresponding to UC and CF deployments usually cross, and, generally, UC approach outperforms the CF approach for the vast majority of users. This means that, excluding a small percentage of MSs with bad channel conditions, the UC approach is beneficial, probably due to the fact

that each AP uses its power to transmit to nearby MSs and avoids wasting power in order to transmit to far MSs.

- The crossing does not take place in Fig. 5.3, wherein the performance corresponding to the maximum sum-rate power allocation policy is reported. This behavior can be explained by noticing that when power allocation is optimized, the CF approach provides a much greater flexibility than the UC case. Indeed, in the UC case, MN downlink transmit powers are optimized, forcing to zero the remaining $M(K-N)$ transmit powers; in the CF case, instead, the system achievable sum-rate over MK transmit powers is optimized, so the solution space is much larger, and it is possible to note that from this point of view the UC approach can be seen as a special case of the CF one[19], and this explain why for this case CF outperforms the UC approach.

- From Fig. 5.4, then, it is seen that, as expected the CDF curves are steeper, since the power allocation maximizing the minimum-rate introduces fairness among the MSs, so there are no large disparities across the rates achieved by each MS.

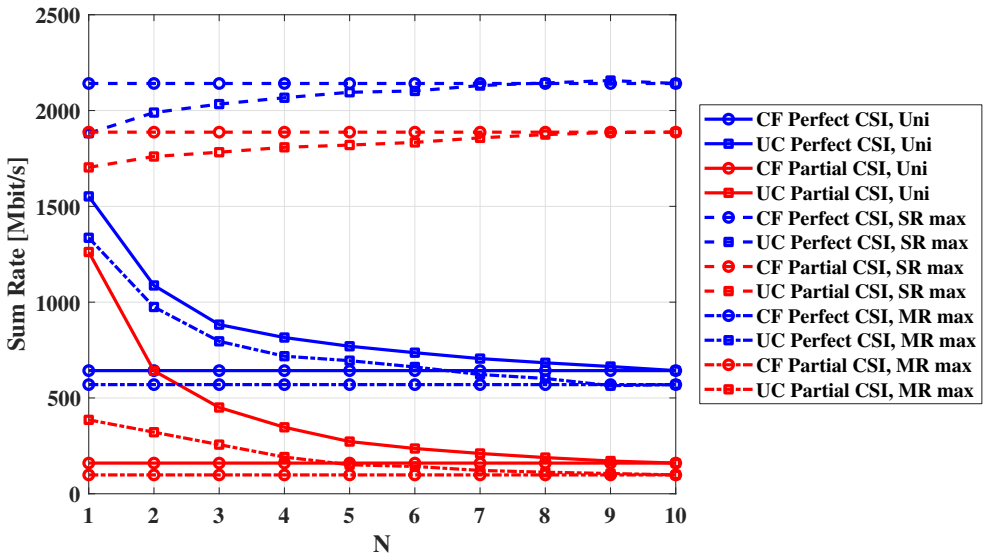

FIGURE 5.5: Sum-rate of the system in downlink versus N. Parameters: $M = 60$, $K = 10$, and $\tau_p = 16$.

Fig. 5.5 shows the average achievable sum-rate of the system in downlink versus N; assuming $M = 60$, $K = 10$, and $\tau_p = 16$. From this figure, it can be noted that, again, focusing on the sum-rate maximizing power allocation, the CF approach outperforms the UC one, whereas, if the power allocation maximizing

[19] This statement should not erroneously lead to the conclusion that the CF deployment is more general and provides better performance of the UC deployment. Indeed, while the statement only applies to downlink, it should also be considered that the CF deployment is not scalable and, also, the power allocation routine requires many iterations to converge given the larger number (with respect to UC approach) of variables to be optimized.

the minimum-rate or the uniform power control allocations are considered, the UC achieves generally better performance, in terms of average achievable sum-rate, than the CF approaches.

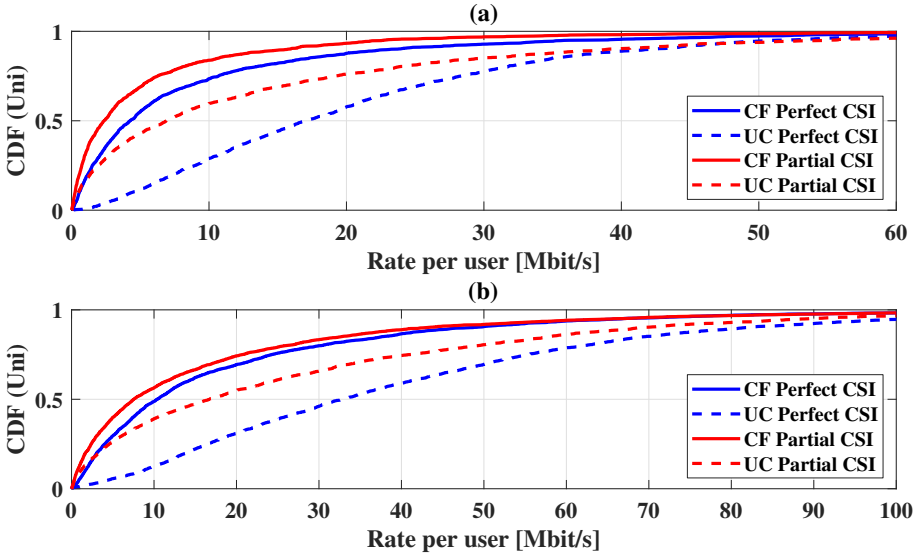

FIGURE 5.6: CDF of rate per user in uplink with uniform power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

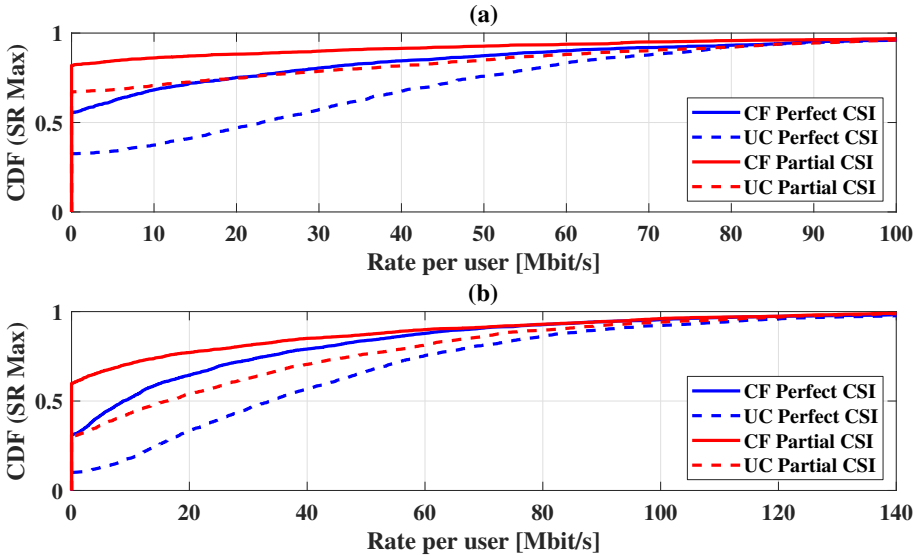

FIGURE 5.7: CDF of rate per user in uplink with sum-rate maximizing power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

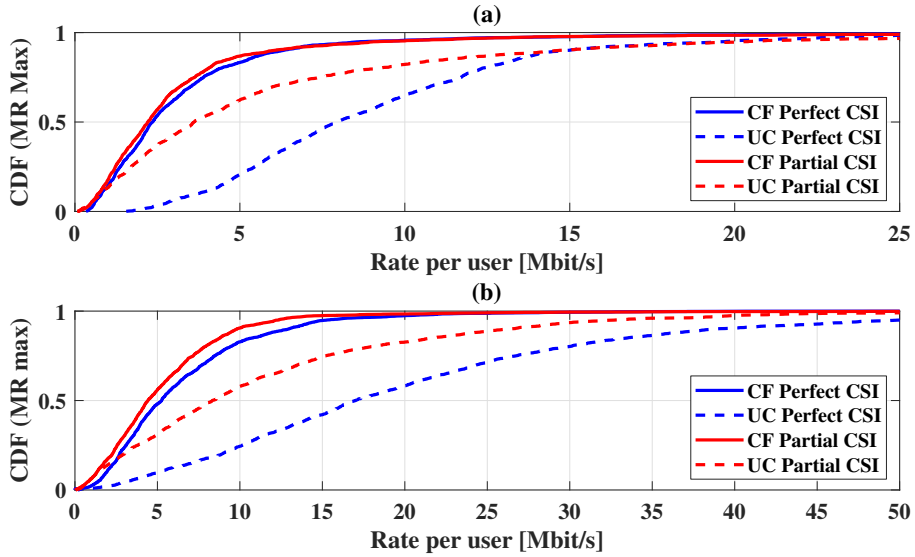

FIGURE 5.8: CDF of rate per user in uplink with minimum-rate maximizing power allocation for a high density scenario in subfigure (a) and for a low density scenario in subfigure (b). Parameters: (a) $M = 80$, $K = 15$, $N = 6$, and $\tau_p = 16$; (b) $M = 50$, $K = 5$, $N = 2$, and $\tau_p = 8$.

With regard to the uplink, in Fig. 5.6 the CDFs of the rate per user for the CF and UC approaches for the case in which uniform power allocation is used are reported. Fig. 5.7 shows the CDFs of the rate per user in uplink for the CF and UC approaches for the case in which sum-rate maximizing power allocation is used and Fig. 5.8 reports the CDFs of the rate per user in uplink for the CF and UC approaches for the case in which minimum-rate maximizing power allocation is used. In Figs. 5.6-5.8 the results of the rate per user in the high and low density scenarios are compared; the parameters used here are the same as in Figs. 5.2-5.4.

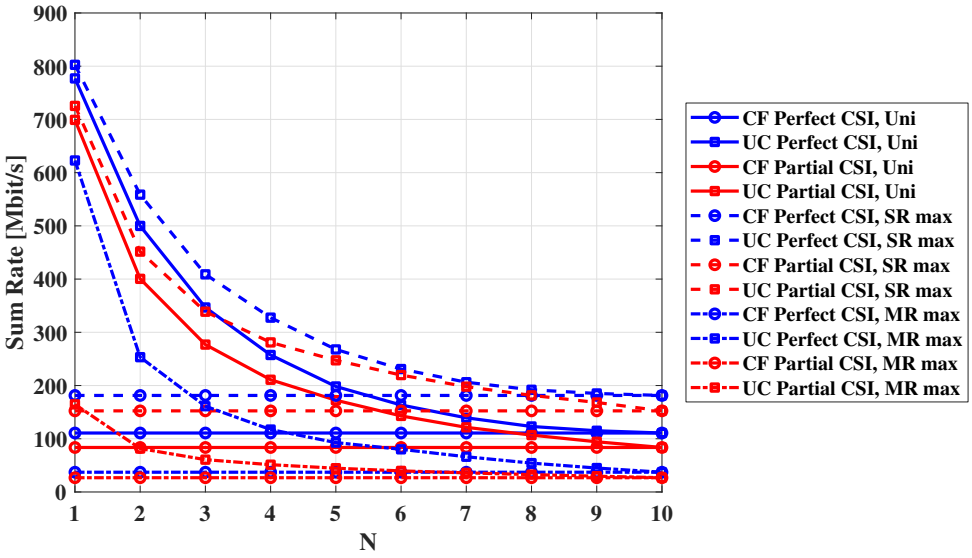

FIGURE 5.9: Sum-rate of the system in uplink versus N. Parameters: $M = 60$, $K = 10$, and $\tau_p = 16$.

Fig. 5.9 shows the system uplink sum-rate versus N, the number of MSs served by each AP, when the number of users is $K = 10$. Inspecting these figures, it can seen that, differently from the downlink, the UC approach outperforms the CF one in all the cases of uniform power allocation and power control strategies, and both in the cases of high and low density scenario. In particular, there are situations in which the UC approach guarantees many-fold improvements with respect to the CF strategy. This behavior can be explained by noting that, for the uplink, the UC can be no longer regarded as a special case of the CF configuration. In the uplink, the CF strategy requires that APs participate to the decoding of far MSs, and this adds a lot of additional noise to the decision statistic that ultimately endanger performance. In the UC approach, instead, each MS is decoded only by nearby APs, that can also rely on good channel estimates, and this helps in considerably increasing the system performance.

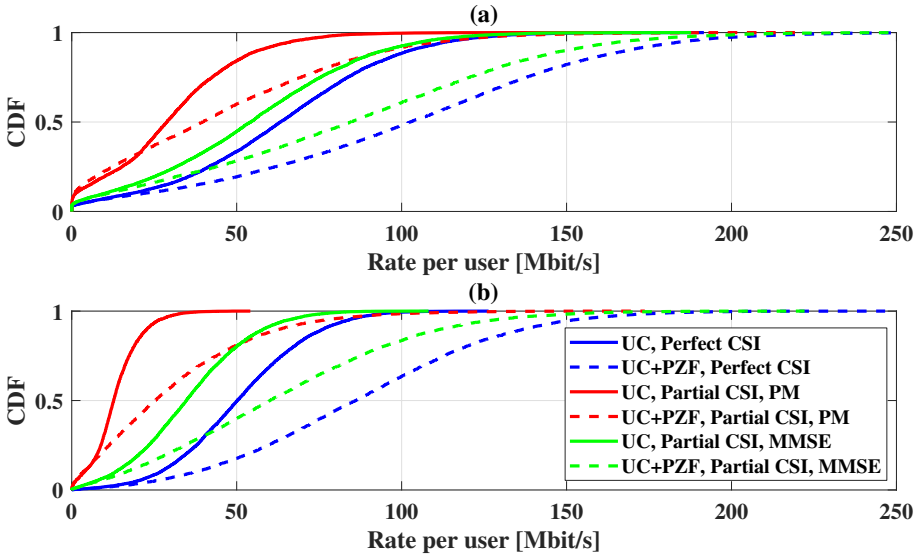

FIGURE 5.10: CDF of rate per user in downlink. Parameters: $M = 100$, $K = 20$, $\tau_p = 16$, in subfigure (a), $N = 2$, and in subfigure (b), $N = 5$.

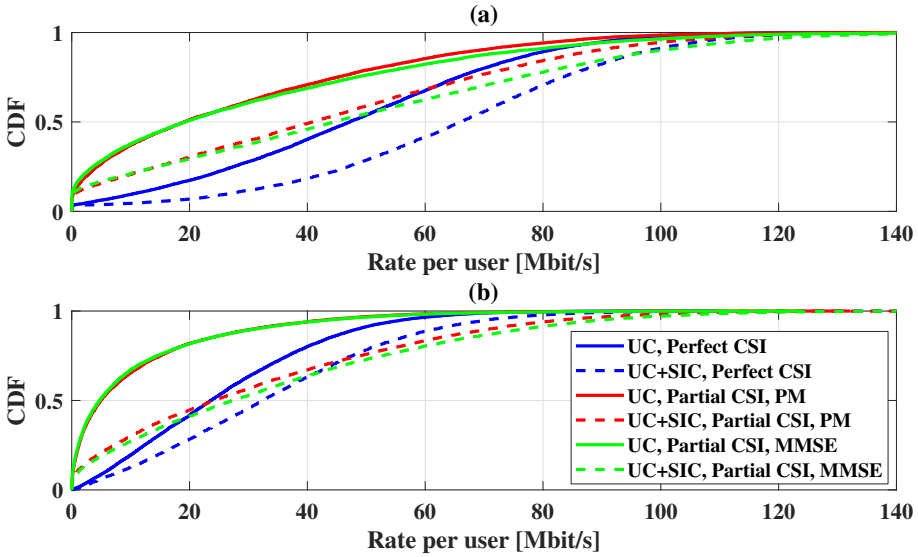

FIGURE 5.11: CDF of rate per user in uplink. Parameters: $M = 100$, $K = 20$, $\tau_p = 16$, in subfigure (a), $N = 2$, and in subfigure (b), $N = 5$.

Regarding the simulation setup for the advanced signal processing tecniques explained in Section 5.6, $P = 1$ is assumed and, for the PZF beamforming in the downlink, $S = 1$, so, in the design of the beamformer at one AP for a generic user, the intereference to the user received at the AP with the highest power is canceled. For the channel estimation, randomly generated binary antipodal pilot sequences are used with length $\tau_p = 16$. In the following figures, "UC" denotes the strategies of CM pre-coding on the downlink, in Sections 5.3.2, and MR post-coding in the uplink, in Section 5.3.3, and "UC+PZF" and "UC+SIC" the ones proposed in Section 5.6.2 for the downlink and in Section 5.6.3 for the

uplink, respectively. In the following figures uniform power allocation is assumed. In Figs. 5.10 and 5.11 the CDFs of the rate per user in downlink and uplink, respectively, are reported, with $N = 2$ and $N = 5$, in the cases of perfect CSI, with the PM channel estimation reported in Section 5.3.1 and with the linear MMSE channel estimation detailed in Section 5.6.1. Inspecting the figures, it is possible to note that both the PZF and the SIC improve the performance of the UC for the vast majority of the MSs in the system. As an example, considering subfigures labeled as (a), the median rate in Fig. 5.10, for the case of MMSE channel estimation, increases from 54 Mbit/s to 84 Mb/s, with a $+55\%$ gain, and in Fig. 5.11 it increases from 20 Mbit/s to 40 Mbit/s, with a $+100\%$ gain. Regarding the AP-MS association rule, numerical evidence shows that having $N = 2$ provides better performance than $N = 5$. Finally, results also show that there is a considerable gap between the case of perfect CSI and the cases in which channel must be estimated. This gap can be explained by noting that pilot sequences are correlated since it is here assumed the use of random pilots with length $\tau_p = 16$ for a number of users $K = 20 > \tau_p$.

Chapter 6

Conclusions and future developments

6.1 Summary of the results

The major topic of this Ph.D. thesis was massive MIMO technology for 5G and beyond-5G wireless networks. In particular, massive MIMO has been studied in conjunction with two other current research topics for the future wireless networks: the mm-Wave frequencies and the distributed antenna systems.
Chapter 1 has given a brief overview on the requirements of the future wireless systems and it has discussed the contribution of the massive MIMO technology to the development of 5G and beyond-5G wireless networks.
Chapter 2 has briefly explained some of the mathematical tools used in the current literature in order to study the performance of the massive MIMO systems and it has given a comparison between different techniques currently used to evaluate the spectral efficiency bounds of such a system.
Chapter 3 has outlined a critical comparison between massive MIMO systems at mm-Waves and at μ-Waves. Six key differences have been outlined, and their implications on the transceiver architecture and on the attainable performance have been discussed and validated also through computer simulations. Among the discussed differences, the most disruptive one was the fact that MIMO systems may be doubly massive at mm-Waves. In this chapter, it has been shown that the use of large-scale antenna arrays has not an as beneficial impact on the system multiplexing capabilities as it has at μ-Wave frequencies. Additionally, the availability of doubly massive MIMO wireless links enables the generation of very narrow beams, resulting in reduced co-channel interference to other users using the same time-frequency resources. Another key advantage of doubly massive MIMO systems at mm-Waves is the fact that the computational complexity of channel estimation weakly depends on the number of antennas, especially for the case in which analog (beam-steering) beamforming strategies are used. While massive MIMO at μ-Wave frequencies is gradually entering 3GPP standards, mm-Waves and in particular massive mm-Wave MIMO systems are still under heavy investigation, both in academia and industry.

Chapter 4 has provided an analysis of a multiuser doubly massive MIMO system operating at mm-Wave frequencies and with several FD and low-complexity beamforming architectures. The obtained results have revealed that, using some of the most recent available data on the energy consumption of transceiver components, FD architectures were superior not only in terms of achievable rate, but also in terms of energy efficiency. In particular, among FD implementations, the PZF-FD architecture has been shown to provide the best performance, while, among the lower complexity implementations, the AN structure can be considered for its extremely low complexity. A detailed analytical study of some beamforming structures in the large number of antennas regime has also been provided, and results have been shown proving the accuracy of the found approximations. Of course the provided results and the relative ranking among the considered structures in terms of energy efficiency is likely to change in the future as technology progresses and devices with reduced power consumption appear on the scene, even though it may be expected that in the long run FD architectures will be fully competitive, in terms of hardware complexity and energy consumption, with HY alternatives.
Chapter 5 has focused on the recently introduced CF massive MIMO architecture. First of all, it has extended the CF approach when the APs and the MSs are equipped with multiple antennas, and it has proposed the use of a channel-inverting beamforming scheme that does not require channel estimation at the MSs. Then, it has compared the CF architecture with the UC approach wherein each AP only decodes a pre-assigned number of MSs. In this chapter two power allocation strategies for the uplink and downlink have been proposed, both for the CF and the UC cases. The first one is a sum-rate maximizing power allocation strategy, aimed at maximizing performance of the system in terms of overall data-rate, while the second one is a minimum-rate maximizing power allocation, aimed at maximizing performance of the system in terms of fairness. Additionally, in this chapter advanced signal processing techniques have been proposed in order to improve the performance of the UC architecture, in particular a local PZF downlink pre-coding at the APs and a SIC procedure for data detection on the uplink have been detailed. Interestingly, both the SIC processing and the downlink PZF pre-coding can be implemented with information available locally at each AP, while, additionally, no channel estimation is required at the MSs. Regarding the numerical results, the performance of the power allocation strategies here proposed have been compared with the case of uniform power allocation and the performance of the advanced signal processing techniques have been studied. Results have shown that the UC approach generally outperforms the CF one, especially on the uplink. The results of the advanced signal processing techniques have shown that the proposed algorithms are capable of providing remarkable performance improvements with respect to the simple CM pre-coding on the downlink and to the MR post-coding in the uplink.

6.2 Future developments

The topic of massive MIMO systems for the future wireless networks is continuously evolving, both in the industry and in academia. During the writing of this thesis, surely new applications of this technology have been conceived and will be studied in the next future. Some future developments of massive MIMO technology are given in the following.

As emphasized in several points of this thesis, massive MIMO technology can offer enhanced broadband services in the future, and more. 5G networks are expected to support a large variety of wireless services in areas ranging from infotainment to healthcare, smart homes and cities, manufacturing, and many others. Massive MIMO technology can be tailored to support a massive number of Machine Type Communication (MTC) devices. Also, it is an excellent candidate to realize ultra-reliable low-latency communications (URLLC) as it can establish very robust physical links [98].

The conjunction of massive MIMO with mm-Wave frequencies allows the consideration of doubly massive MIMO systems. The interest in this topic can be justified by the high number of use cases that can be implemented to exploit the availability of large bandwidths and consequently high data rates available at mm-Wave [99]. Realistic use-cases can be envisioned already with the current technology, for example: the use of large arrays for the vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, in order to increase also the reliability of the communication using the antenna gain to obtain a higher received SNR; the use of doubly massive MIMO systems for the Fixed Broadband Wireless Access (FBWA); the use of large antenna arrays at both sides of the communications in order to implement the wireless backhaul in a cellular network. Another interesting research topic regarding the use of doubly massive MIMO systems should be the study of modulation formats alternative to the OFDM for these systems. The high peak-to-average power ratio (PAPR) of the OFDM works against the advantages of using large antenna arrays, and can impede good downlink performance. Recent studies indicates that single-carrier modulations (SCMs) at mm-Wave can achieve good performance in presence of non linear PA [39] for the single user communication. These studies can be extended to the case of multiuser systems, in presence of imperfect CSI and with waterfilling-like power control.

The conjunction of massive MIMO with distributed antenna systems gives some relevant research topics that can be investigated. Among others, the following are pointed out: the consideration of the such a system at mm-Wave frequencies (preliminary results on this are reported in [100], [101]); the coupling of distributed massive MIMO architectures with multiple access schemes such as the non-orthogonal multiple access (NOMA); the introduction of unmanned aerial vehicles (UAVs) in a distributed massive MIMO system, as additional APs or as MSs, as recently discussed for a co-located massive MIMO system in [102]

and considering the co-existence of a distributed massive MIMO system with a radar system on the same frequency band, as preliminarly reported in [103] considering a co-located massive MIMO system.

Bibliography

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

[2] D. H. Ring, "Mobile telephony-wide area coverage", *Bell Laboratories Technical Memorandum*, 1947.

[3] A. Lozano and A. M. Tulino, "Capacity of multiple-transmit multiple-receive antenna architectures", *IEEE Transactions on Information Theory*, vol. 48, no. 12, pp. 3117–3128, Dec. 2002.

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

[5] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

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

[7] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G", *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[8] N. Jindal and A. Lozano, "A unified treatment of optimum pilot overhead in multipath fading channels", *IEEE Transactions on Communications*, vol. 58, no. 10, pp. 2939–2948, Sep. 2010.

[9] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?", *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[10] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?", *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.

[11] E. Björnson, J. Hoydis, and L. Sanguinetti, "Pilot contamination is not a fundamental asymptotic limitation in massive MIMO", in *Proc. of 2017 IEEE International Conference on Communications (ICC)*, May 2017, pp. 1–6.

[12] Y. Zeng, R. Zhang, and Z. N. Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO", in *Proc. of 2013 IEEE/CIC International Conference on Communications in China (ICCC)*, Aug. 2013, pp. 454–459.

[13] Y. Zeng, R. Zhang, and Z. N. Chen, "Electromagnetic lens-focusing antenna enabled massive MIMO: Performance improvement and cost reduction", *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1194–1206, Jun. 2014.

[14] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology", *IEEE Communications Magazine*, vol. 51, no. 6, pp. 172–179, 2013.

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

[16] R. N. Almesaeed, A. S. Ameen, E. Mellios, A. Doufexi, and A. Nix, "3D channel models: Principles, characteristics, and system implications", *IEEE Communications Magazine*, vol. 55, no. 4, pp. 152–159, Apr. 2017.

[17] A. Adhikary, E. A. Safadi, and G. Caire, "Massive MIMO and inter-tier interference coordination", in *2014 Information Theory and Applications Workshop (ITA)*, Feb. 2014, pp. 1–10.

[18] J. Chen and V. K. N. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks", *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1230–1238, Jun. 2014.

[19] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?", *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[20] 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 Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[21] 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.

[22] S. M. Kay, *Fundamentals of statistical signal processing, volume 1: Estimation theory*. 1998.

[23] G. Caire, "On the ergodic rate lower bounds with applications to massive MIMO", *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3258–3268, May 2018.

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

[25] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems", *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[26] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!", *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[27] *Huawei and DOCOMO Conduct World's First 5G Large Scale Field Trial in the 4.5 GHz Band*, [Online]. Available: `http://www.huawei.com/en/news/2016/11/World-First-5G-Large-Scale-Field-Trial`, Nov. 2016.

[28] *Ericsson 5G field trial gear achieves peak downlink throughput over 25 Gbps with MU-MIMO*, [Online]. Available: `https://www.ericsson.com/news/1987136`, Feb. 2016.

[29] S. Buzzi and C. D'Andrea, "Doubly massive mmwave MIMO systems: Using very large antenna arrays at both transmitter and receiver", in *Proc. of 2016 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016, pp. 1–6.

[30] E. Björnson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences", *ArXiv preprint arXiv:1803.11023*, Mar. 2018.

[31] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless personal communications*, vol. 6, no. 3, pp. 311–335, 1998.

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

[33] *5G Channel Model for bands up to 100 GHz*, `http://www.5gworkshops.com/5GCM.html`, 2015.

[34] 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 Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[35] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells", *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[36] A. Tang, J. Sun, and K. Gong, "Mobile propagation loss with a low base station antenna for NLOS street microcells in urban area", in *Proc. of Vehicular Technology Conference (VTC Spring), 2001 IEEE 53rd*, IEEE, vol. 1, May 2001, pp. 333–336.

[37] 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 Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, Sep. 2015.

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

[39] S. Buzzi, C. D'Andrea, T. Foggi, A. Ugolini, and G. Colavolpe, "Single-carrier modulation versus OFDM for millimeter-wave wireless MIMO", *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1335–1348, Mar. 2018.

[40] J. Lee, G.-T. Gil, and Y. H. Lee, "Exploiting spatial sparsity for estimating channels of hybrid MIMO systems in millimeter wave communications", in *Proc. of 2014 IEEE Global Communications Conference (GLOBECOM)*, IEEE, 2014, pp. 3326–3331.

[41] S. Buzzi and C. D'Andrea, "Are mmwave low-complexity beamforming structures energy-efficient? analysis of the downlink MU-MIMO", in *Proc. of 2016 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–6.

[42] S. Buzzi and C. D'Andrea, "Energy efficiency and asymptotic performance evaluation of beamforming structures in doubly massive MIMO mmwave systems", *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 2, pp. 385–396, Jun. 2018.

[43] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC", *ArXiv preprint arXiv:1404.7736*, Apr. 2014.

[44] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?", *IEEE Access*, vol. 4, pp. 247 –267, Mar. 2016.

[45] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays", *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 501–513, Jan. 2016.

[46] M. D. Migliore, "On the role of the number of degrees of freedom of the field in MIMO channels", *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 2, pp. 620–628, Feb. 2006.

[47] S. Haghighatshoar and G. Caire, "Massive MIMO channel subspace estimation from low-dimensional projections", *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 303–318, Jan. 2017.

[48] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems", *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 528–542, Apr. 2016.

[49] S. Buzzi and C. D'Andrea, "Subspace tracking algorithms for millimeter wave MIMO channel estimation with hybrid beamforming", in *Proc. of WSA 2017; 21th International ITG Workshop on Smart Antennas*, Mar. 2017, pp. 1–6.

[50] K. Abed-Meraim, S. Attallah, A. Chkeif, and Y. Hua, "Orthogonal Oja algorithm", *IEEE Signal Processing Letters*, vol. 7, no. 5, pp. 116–119, May 2000.

[51] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals", *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[52] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination", *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[53] N. Krishnan, R. D. Yates, and N. B. Mandayam, "Uplink linear receivers for multi-cell multiuser MIMO with pilot contamination: Large system analysis", *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4360–4373, Aug. 2014.

[54] S. A. R. Naqvi, S. A. Hassan, and Z. ul Mulk, "Pilot reuse and sum rate analysis of mmwave and UHF-based massive MIMO systems", in *Proc. of Vehicular Technology Conference (VTC Spring), 2016 IEEE 83rd*, IEEE, 2016, pp. 1–5.

[55] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems", *IEEE Communications Magazine*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[56] X. Zhang, Z. Lv, and W. Wang, "Performance analysis of multiuser diversity in MIMO systems with antenna selection", *IEEE Transactions on Wireless Communications*, vol. 7, no. 1, pp. 15–21, Jan. 2008.

[57] X. Ge, R. Zi, H. Wang, J. Zhang, and M. Jo, "Multi-user massive MIMO communication systems based on irregular antenna arrays", *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5287–5301, Aug. 2016.

[58] S. Buzzi, "Doubly-massive MIMO systems at mmWave frequencies: Opportunities and research challenges", in *IEEE WCNC'2016 Workshop on Green and Sustainable 5G Wireless Networks*, keynote talk, Doha, Qatar, Apr. 2016.

[59] S.Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G", *IEEE Communications Magazine*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[60] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems", *IEEE Transactions on Communications*, Apr. 2016.

[61] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation and hybrid combining for mmwave: Phase shifters or switches?", in *Information Theory and Applications Workshop (ITA), 2015*, IEEE, 2015, pp. 90–97.

[62] A. Alkhateeb, Y.-H. Nam, J. Zhang, and R. W. Heath, "Massive MIMO combining with switches", *IEEE Wireless Communications Letters*, vol. 5, no. 3, pp. 232–235, Jan. 2016.

[63] S. Buzzi, C.-L. I, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead", *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, Apr. 2016.

[64] R. Zi, X. Ge, J. Thompson, C. X. Wang, H. Wang, and T. Han, "Energy efficiency optimization of 5G radio frequency chain systems", *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 758–771, Apr. 2016.

[65] K. Roth, J. Garcia, J. Munir, M. Faerber, and J. A. Nossek, "Channel capacity comparison of different system concepts for mmwave", in *Proc. of WSA 2016; 20th International ITG Workshop on Smart Antennas*, Mar. 2016, pp. 1–7.

[66] S. Haghighatshoar and G. Caire, "Enhancing the estimation of mm-Wave large array channels by exploiting spatio-temporal correlation and sparse scattering", in *Proc. of WSA 2016; 20th International ITG Workshop on Smart Antennas*, Mar. 2016, pp. 1–7.

[67] F. Negro, S. P. Shenoy, I. Ghauri, and D. T. Slock, "On the MIMO interference channel", *Proc. of Information Theory and Applications Workshop 2010*, pp. 1–9, Feb. 2010.

[68] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory", *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2015.

[69] A. V. den Bosch, M. A. F. Borremans, M. S. J. Steyaert, and W. Sansen, "A 10-bit 1-GSample/s Nyquist current-steering CMOS D/A converter", *IEEE Journal of Solid-State Circuits*, vol. 36, no. 3, pp. 315–323, Mar. 2001.

[70] L. Kong, "Energy-efficient 60GHz phased-array design for multi-Gb/s communication systems", PhD thesis, EEECS Department University of California, Berkeley, Dec. 2014.

[71] Y.-Y. Lee and C.-H.W.Y.-H. Huang, "A hybrid RF/Baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems", *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 305 –317, Nov. 2014.

[72] Y. Yu, P. G. M. Baltus, A. de Graauw, E. van der Heijden, C. S. Vaucher, and A. H. M. van Roermund, "A 60 GHz phase shifter integrated with LNA and PA in 65 nm CMOS for phased array systems", *IEEE Journal of Solid-State Circuits*, vol. 45, no. 9, pp. 1697–1709, Sep. 2010.

[73] H. Ghauch, M. Bengtsson, T. Kim, and M. Skoglund, "Subspace estimation and decomposition for hybrid analog-digital millimetre-wave MIMO systems", in *Proc. of 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, Jun. 2015, pp. 395–399.

[74] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Designing multi-user MIMO for energy efficiency: When is massive MIMO the answer?", in *Proc. of 2014 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2014, pp. 242–247.

[75] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?", *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, Feb. 2015.

[76] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination", *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 1–14, Jan. 2015.

[77] S. Zhou, M. Zhao, X. Xu, J. Wang, and Y. Yao, "Distributed wireless communication system: A new architecture for future public wireless access", *IEEE Communications Magazine*, vol. 41, no. 3, pp. 108–113, Mar. 2003.

[78] K. T. Truong and R. W. Heath, "The viability of distributed antennas for massive MIMO systems", in *Proc. of 2013 Asilomar Conference on Signals, Systems and Computers*, Nov. 2013, pp. 1318–1323.

[79] W. Feng, Y. Wang, N. Ge, J. Lu, and J. Zhang, "Virtual MIMO in multi-cell distributed antenna systems: Coordinated transmissions with large-scale CSIT", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2067–2081, Oct. 2013.

[80] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone", in *Proc. of 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jul. 2015, pp. 201–205.

[81] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?", in *Proc. of 2016 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016, pp. 1–7.

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

[83] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design", *IEEE Communications Letters*, vol. 21, no. 8, pp. 1871–1874, Aug. 2017.

[84] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers", in *Proc. of 2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov. 2016, pp. 203–207.

[85] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO", *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 1, pp. 25–39, Mar. 2018.

[86] Q. Huang and A. Burr, "Compute-and-forward in cell-free massive MIMO: Great performance with low backhaul load", in *Proc. of 2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2017, pp. 601–606.

[87] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free massive MIMO with limited backhaul", *ArXiv preprint arXiv:1801.10190*, Jan. 2018.

[88] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach", *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 706–709, Dec. 2017.

[89] S. Buzzi and C. D'Andrea, "User-centric communications versus cell-free massive MIMO for 5G cellular networks", in *Proc. of WSA 2017; 21th International ITG Workshop on Smart Antennas*, Mar. 2017, pp. 1–6.

[90] S. Buzzi and A. Zappone, "Downlink power control in user-centric and cell-free massive MIMO wireless networks", in *Proc. of 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–6.

[91] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems", in *Proc. of 2010 Future Network Mobile Summit*, Jun. 2010, pp. 1–10.

[92] Z. Wang, E. K. Tameh, and A. R. Nix, "Joint shadowing process in urban peer-to-peer radio channels", *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 52–64, Jan. 2008.

[93] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization", *SIAM Journal on Optimization*, vol. 23, no. 2, Jun. 2013.

[94] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[95] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for non-convex mathematical programs", *Operations Research*, vol. 26, no. 4, pp. 681–683, Aug. 1978.

[96] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to non-convex truss topology design problems", *Journal of Global Optimization*, vol. 47, no. 1, Jul. 2009.

[97] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization*, 2015.

[98] E. G. Larsson and L. Van der Perre, "Massive MIMO for 5G", *IEEE 5G Tech Focus*, vol. 1, no. 1, Mar. 2017.

[99] S. Buzzi, "The case of doubly massive MIMO systems at millimeter wave", in *IEEE Communication Theory Workshop*, keynote talk, Miramar Beach, Florida, USA, May 2018.

[100] M. Alonzo and S. Buzzi, "Cell-free and user-centric massive MIMO at millimeter wave frequencies", in *Proc. of 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Oct. 2017, pp. 1–5.

[101] M. Alonzo, S. Buzzi, and A. Zappone, "Energy-efficient downlink power control in mmwave cell-free and user-centric massive MIMO", in *Proc. of 2018 IEEE 1st 5G World Forum (5GWF'18)*, Jul. 2018.

[102] G. Geraci, A. Garcia-Rodriguez, L. G. Giordano, D. López-Pérez, and E. Björnson, "Understanding UAV cellular communications: From existing networks to massive MIMO", *ArXiv preprint arXiv:1804.08489*, Apr. 2018.

[103] S. Buzzi, M. Lops, C. D'Andrea, and C. D'Elia, "Co-existence between a radar system and a massive MIMO wireless cellular system", in *Proc. of 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2018, pp. 1–5.

Acknowledgements

These last pages of the thesis are dedicated to the acknowledgements, in these few lines I want to thank all the people who in various ways have supported (and endured) me in these three years of Ph.D. studies. This is not a formality, but it comes from the wish to express my gratitude to all the people who contributed to my human and professional growth during these three years. First of all I want to thank my research supervisors, who always have guided me with kindness and experience, Prof. Stefano Buzzi for his constant human and scientific support during these three years, and Dr. Alessio Zappone for his valuable suggestions and for his kind support always demonstrated. I thank Prof. Angel Lozano, research supervisor during my period of study at the Universitat "Pompeu Fabra" in Barcelona (Spain), for his guidance and suggestions on my research topics. I would to thank the colleagues of the Wireless Communications (WiCom) Research Group of the Universitat "Pompeu Fabra" in Barcelona for having kindly welcome me and for our interesting discussions on various research topics that have contributed to my human and professional growth. I want to thank all the people that I met during the courses, seminars and Ph.D. schools, that I attended in the University of Cassino and Southern Lazio and in other universities, that have contributed to fuel my interest in the research topics on 5G and beyond-5G networks. I am particularly grateful to all the people who believed in me and who pushed me to always do my best, my parents, my sister Antonella and my brother Gianpaolo, my grandparents, my "true friends" Loredana, Ida and Federica and my "angels" that I felt always with me. A special thanks to Luca, who has always been close to me in the most important goals of my life and he has always supported me with his unconditional love. Finally, I want to thank all the readers of this thesis, in the hope that they have found in this contribution something interesting and useful in the research on the massive MIMO technologies for future cellular networks.

Ringraziamenti (in italian)

Queste ultime pagine della tesi sono dedicate ai ringraziamenti, in queste poche righe voglio ringraziare tutte le persone che in vari modi mi hanno supportata (e sopportata) in questi tre anni di dottorato. Questa non è una formalità, ma nasce dal mio desiderio di ringraziare tutte quelle persone che hanno contribuito alla mia crescita umana e professionale durante questi tre anni di dottorato. Innanzitutto vorrei ringraziare i miei supervisori, che mi hanno guidata sempre con gentilezza ed esperienza, il Prof. Stefano Buzzi per il suo costante supporto umano e scientifico durante tutto il mio percorso di studio e di vita in questi tre anni, e il Dr. Alessio Zappone per i suoi preziosi consigli e per il suo gentile supporto sempre dimostratomi. Ringrazio il Prof. Angel Lozano, supervisore durante il mio periodo di studio presso l' Universitat "Pompeu Fabra" di Barcellona (Spagna), per la sua guida e i suoi suggerimenti sui miei argomenti di ricerca. Ringrazio i colleghi del gruppo di ricerca Wireless Communications (WiCom) Research Group dell'Universitat "Pompeu Fabra" di Barcellona per avermi fatta sentire a casa e per le interessanti discussioni su diversi argomenti di ricerca che hanno contribuito alla mia crescita umana e professionale. Voglio ringraziare tutte le persone che ho incontrato durante i corsi, i seminari e le scuole di dottorato, frequentati presso l'Università di Cassino e del Lazio Meridionale e presso altre università, che hanno contribuito ad alimentare il mio interesse negli argomenti di ricerca sulle reti 5G and beyond-5G. Sono particolarmente grata a tutte le persone che hanno creduto in me e che mi hanno spinta a fare sempre del mio meglio, ai miei genitori, a mia sorella Antonella e a mio fratello Gianpaolo, ai miei nonni, alle mie "amiche vere" Loredana, Ida e Federica ed ai miei "angeli" che sento sempre con me. Un ringraziamento particolare a Luca, che mi è sempre stato vicino nei traguardi più importanti della mia vita e mi ha sempre supportata con il suo amore incondizionato. Infine, voglio ringraziare tutti i lettori di questa tesi, nella speranza che abbiano trovato in tale contributo qualcosa di interessante e di utile nell'ambito della ricerca sulle tecnologie del massive MIMO per le future reti cellulari.