UNIVERSITÀ DEGLI STUDI DI CASSINO E DEL LAZIO MERIDIONALE


DIPARTIMENTO DI INGEGNERIA ELETTRICA E DELL'INFORMAZIONE

---

# Distributed MIMO Systems 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*

June 2021

*RESEARCH SUPERVISORS:*
Prof. Stefano BUZZI
Prof. Alessio ZAPPONE

*AUTHOR:*
Mario Alonzo

# Abstract

Massive Multiple-Input-Multiple-Output (MIMO) is a recent technlogy that will be exploited for 5G and beyond-5G wireless network due to the constraints given by the future wireless networks, such as low latency and high spectral efficiency. In this thesis, MIMO systems have been taken into account in order to study two different network architectures. The former is called Cell-Free (CF), and it has been studied at millimeter Wave (mmWave) and microwave frequencies, and the latter is called Distributed Multiple Input Multiple Output (D-MIMO) for factory automation. The first chapter of this thesis gives an overview of massive MIMO, so why there is the need to exploit this technology and gives some mathematical concept. In the second chapter the CF at mmwave frequencies has been studied. The CF is a recent network architecture, in order to alleviate the cell-edge problem and thus increase the system performance of unlucky users that happen to be located very far from their serving Access Point (AP). In this architecture a large number of distributed APs, connected to a central processing unit (CPU), simultaneously and jointly serve a much smaller number of mobile Station (MS) or users. Both APs and users are equipped with multiple antennas. Then, it has been analyzed an architecture that generalizes the CF, the so called User-centric (UC), where each AP has to serve only a limited number of users. A power control algorithm has been introduced by resorting a method called successive lower bound maximization, aimed at maximizing the sum-rate and the energy efficiency. At mmwave, a lot of antennas can be employed, this means that there is the need of using hybrid architecture at each AP in order to reduce complexity and cost by using a small number of radio frequency (RF) chains. With CF or UC, channel estimation and beamforming are locally evaluated, reducing the traffic load on the backhaul network. So, a comparison between a fully-digital (FD) and hybrid (HY) architecture will be shown. What it is possible to anticipate is that the FD architecture provides better performances than the hybrid one. In the numerical results, the performances in term of energy efficiency and sum-rate on Downlink and Uplink, with uniform and optimal

power allocation and with a fully digital and hybrid architectures will be addressed. Then, this thesis also focuses on the comparison between D-MIMO and CF architectures for factory automation, at microwave frequencies. In this case, communications between actuators (ACs) and APs inside an industrial scenario is considered by adopting those different communication systems. Then, different transmission modes are taken into account, Joint transmission joint transmission (JT), Cell-Free transmission (CFT), single AP transmission (SAT), and User-centric transmission (UCT). In SAT mode each AC is served by only one AP. Even for this scenario a power control rule has been taken into account. In the end, in numerical section, it has been shown the performances in terms of SINR and achievable rate, evaluated with the finite block length capacity (FBLC) formula, when different transmission modes and beamformers are employed, and moreover the improvement given by the use of a power control.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

**A/D** analog-to-digital . . . . . 3

**AP** Access Point . . . . . i

**BS** base station . . . . . 7

**CF** Cell-Free . . . . . i

**CFT** Cell-Free transmission . . . . . ii

**CPU** central processing unit . . . . . i

**CSI** channel state information . . . . . 41

**D/A** digital-to-analog . . . . . 3

**D-MIMO** Distributed Multiple Input Multiple Output . . . . . i

**EHF** Extremely High Frequency . . . . . 2

**eMBB** Enhanced Mobile Broadband . . . . . 1

**FBLC** finite block length capacity . . . . . ii

**FD** fully-digital . . . . . i

**FZF** full Zero-forcing . . . . . 7

**HY** hybrid . . . . . i

**JT** joint transmission . . . . . ii

# Mathematical Notation

| | |
|---|---|
| $\mathbf{A}$ | generic matrix |
| $\mathbf{a}$ | generic column vector |
| $\mathcal{A}$ | generic set $\mathcal{A}$ |
| $a$ | generic scalar |
| $a \in A$ | a is a member of a set $\mathcal{A}$ |
| $a \notin A$ | a is not a memeber of a set $\mathcal{A}$ |
| $\mathbb{C}^{N \times M}$ | the set of complex-valued $N \times M$ matrices |
| $\mathbb{R}^{N \times M}$ | the set of real-valued $N \times M$ matrices |
| $\mathbf{I}_N$ | $N \times N$ identity matrix |
| $(A)^T$ | transpose of matrix A |
| $(A)^H$ | conjugate transpose of matrix A |
| $(A)^*$ | conjugate of matrix A |
| $\mathrm{tr}(\mathbf{A})$ | trace of matrix $\mathbf{A}$ |
| $\parallel \mathbf{A} \parallel$ | norm of a matrix $\mathbf{A}$ |
| $\mid \mathbf{A} \mid$ | determinant of matrix $\mathbf{A}$ |
| $\mathbf{A}_{r,:}$ | $r$-th row of matrix A |
| $\mathbf{A}_{:,c}$ | $c$-th column of matrix A |
| $\lvert\mathcal{A}\rvert$ | cardinality of a set $\mathcal{A}$ |
| $\mathbf{x} \sim \mathcal{CN}(\mu, \sigma^2)$ | Complex Gaussian random variable with mean $\mu$ and variance $\sigma^2$ |
| $A \otimes B$ | Kronecker product between two matrices, A and B |
| $\mathrm{card}(\cdot)$ | set's cardinality |
| $a \in \mathcal{S}$ | a belong to a set $\mathcal{S}$ |
| $a \in \mathcal{S}$ | a does not belong to a set $\mathcal{S}$ |
| $\mathrm{vec}(A)$ | vectorized version of matrix A |
| $\nabla$ | gradient operator |

| | |
|---|---|
| $e$ | Euler's number |
| $i$ | imaginary unit of a complex number |
| $\log_a$ | logarithm function with base $a$ |
| $\mathbb{1}[\cdot]$ | indicator function |
| $\max(a, b)$ | maximum between a and b |
| $\max(a, b)$ | minimum between a and b |
| $Q^{-1}(\cdot)$ | inverse Q function |

# Chapter 1

# Introduction

## 1.1 Overview on 5G

5G (and beyond) is the term used to indicate the new wireless networks. The main goals of 5G are:

- increase spectral and area spectral efficiency;
- low latency;
- high reliability and availability;
- support a very huge number of simultaneous connection

This new technology will allow to have different use cases, and each of them belong to a certain of services, such as Enhanced Mobile Broadband (eMBB), Massive machine type communication (mMTC), and Ultra Reliable Low Latency Communications (URLLC) [1]–[3]. In eMBB the goal is to increase the network capacity and the download and upload speed; this will be useful, for example, either for watching video in high definition or for supporting the use of Augmented Reality/Virtual Reality. In mMTC there is the need of providing connections to a large number of different kind of devices that transmit small amounts of traffic, with low power consumption, like sensors used in industrial applications. In URLLC the goal is to reach a very low latency (the target is 1ms), and very high reliability (99.999%). These requirements can be useful for healthcare, transport, industrial automation. So, over

the years, cellular network have been evolving a lot, since in the beginning they were designed only for voice application. All these use cases will be supported by exploiting the so called mmWave frequencies, which permit to use a large bandwidth, in conjunction with antenna array having a lot of antenna elements. Thus, in 5G, massive MIMO, D-MIMO systems and beamforming will play a fundamental role. The thesis focuses on these topic. As it will be shown, the MIMO technology is used to improve the realiability by exploiting the spatial divesity.

As said before, in the future there will be a relevant increase in term of simultaneous connections due to the huge number of devices, so there is the need to use a larger spectrum frequency; this is possible by exploiting the so called mmWave, belonging to the frequency range 30GHz - 300GHz, also known as Extremely High Frequency (EHF). Carrier frequencies at mmWaves allow to have larger bandwidth allocation, which translate directly to higher data transfer and furthermore, to expand the channel bandwidth over 20 MHz used by LTE. At these high frequencies the wavelength is really small, and so the antenna are small too; this means that there the possibility of using many antennas so we have the massive MIMO systems. Thanks to this enhancements, BSs, as well as backhaul links, will be able to handle much greater capacity than LTE system. As said in the introduction, at mmWaves frequencies in addition to this aspects, there are problem regarding the propagation; the pathloss is relevant, because we know higher is the frequency higher are the losses but this problem can be overcame by using large antenna array.

Atmospheric absorption is not a relevant problem over small distances (for small cell at certain frequencies). Very important is the issue regarding penetration through obstacles. The electromagnetic waves at those frequencies are not able to pass through obstacles and the diffraction phenomena is reduced, so LOS condition becomes important in order to have a communication. Large antennas array used in this system, allow to have directional beams, almost as flashlights, and in this way the interference behavior changes completely as well as the sensitivity to misaligned beams. Interference occurs with an on/off behavior, where just the stronger beams interfere[4].

*Link acquisition*, is another challenge at mmWave frequencies, since is tough to establish communication among users and APs, as well as initial access and handoff. In order to find each other, users and APs may need to scan a lots of angular positions where a narrow beam could possible be found, or could be used other strategies.
A way to get over the issue of mmWaves, is to use both microwaves and mmWaves because in this way it possible to have a stable and reliable connection; sporadic interruptions due to mmWaves links would be less frequent[4]. At those frequencies there is the issue related to the hardware, since it is present a relevant power consumption due to the analog-to-digital (A/D) and digital-to-analog (D/A) converters, needed for large bandwidths. Even though large antenna arrays are necessary in order to overtake the pathloss problems, fully digital beamforming for each antenna is unfeasible. As we will see in Chapter 3, a possible solution is to use hybrid beamforming. Let us see very briefly, in this introduction, what is a massive MIMO system. The following concepts, that we are going to explain, are all well described in [4].
A simple MIMO system is essentially a system wherein both transmitter and receiver are equipped with more than one antenna in order to guarantee a better performances. The LTE was developed by using MIMO in which two or four antennas per mobile and up to eight for BSs. In the future standard will be used the massive MIMO system which allows to use at BS a huge number of antennas. This approach offers some benefits:

1. enhancements in spectral efficiency without the need to increase the BS densification;

2. better spatial diversity;

3. simple transmit/receive structures because the quasi-orthogonal nature of the channels between BS and a set of active users.

However, for massive MIMO systems, several challenges must be overcome.
*Pilot contamination*

Pilot transmission can be made orthogonal among same-cell users, to facilitate cleaner channel estimates, but must be reused across cell, otherwise the resources would end up because of pilot transmission. Then, this cause the interference, so called pilot contamination, ant it is does not vanish even if the number of antennas at BS becomes large. However there are methods to reduce or even eliminate this problem.

*Architectural Challenges*

The architecture of BS is quite different; instead of having a few high-power amplifiers feeding a handful of sector of antennas, we would have a lot of tiny antennas fed by correspondingly low-power-amplifiers, this means that each antenna shoud have its own amplifier.

*Full-Dimension MIMO*

Existing BSs are equipped with horizontal arrays, which can only accommodate a limited number of antennas, and which only exploit the azimuth angle. Instead, by using 2D planar array and further exploiting the elevation angle, so called full-dimension MIMO(FD-MIMO), can house more antennas. As a benefit this structure allows to increase the signal power and to reduce interference in neighboring cells.

*Channel models*

With the change of architecture there is the need to change also the channel, since beside azimuth angle, now we are dealing with the elevation angle.

*Coexistence with small cells*

Massive MIMO BSs would have to coexist with small cells, which would not be equipped with massive MIMO. An alternative is to segregate the corresponding transmission in frequency, the large number of excess antennas at massive MIMO BSs may offer the opportunity of reducing interference with relative simplicity.

*Coexistence with mmWaves*

As said before mmWaves require many antennas for beamsteering. The antennas are much smaller at these frequencies and so a large number of antennas can be used for portable devices. In this way it is possible to provide beamforming power gain. The applications at mmWave frequencies would

have to find correct balance between power gain/interference.
In conjuction with the Massive MIMO system, the beamforming will play an important role. Beamforming is a signal processing technique used in order to direct a signal in a precise direction. This can be done by combining elements in an antenna array in such a way that signals at particular angles experience constructive interference while others experience destructive interference. The improvement compared with omnidirectional reception/transmission is known as the directivity of the array.
Also D-MIMO will exploit the MIMO architecture. A D-MIMO is a system where the antennas are geographically distributed among the cells instead of being placed in a single and bigger array. With this system, it is useful to exploit the spatial diversity in order to satisfy the requirements of high spectral efficiency and especially low latency, instead of time and frequency diversity. Time diversity is based on the retrasmission of a data packet, this makes the system robust against the error but not suitable for low latency applications; frequency diversity allows the transmission of the same message signal at different carrier frequencies, and this solution is not suitable for narrowband systems. In the end, spatial diversity is the most used strategy in this kind of applications, since the same information is sent along different spatial path and this increase the spectral efficiency and enhance the system reliability.

## 1.2 Thesis contributions

This thesis focuses on massive MIMO systems at mmWave frequencies and distributed antenna systems for factory automation. Therefore, as it will shown, the novelty of this thesis regards the introduction of a recent network architecture. This architecture, it will be analyzed at mmWave frequecies, in Chapter 3. Then, in Chapter 4, an industrial scenario is considered with the comparison between D-MIMO and CF, at microwave frequencies. Thus, the goal of this thesis is to highlight the potential of the distributed antenna systems, in terms of rate, energy efficiency and SINR, and in different scenarios. Here, we give a brief summary of the thesis:

- **Chapter 3: Cell-Free Massive MIMO at mmWave frequncies**
  CF is a recent network architecture, in order to alleviate the cell-edge problem and thus increase the system performance of unlucky users that happen to be located very far from their serving AP. In this architecture a large number of distributed APs, connected to a CPU, simultaneously and jointly serve a much smaller number of MS or users. Both APs and users are equipped with multiple antennas. In this chapter will be also considered the architecture that generalizes the CF, the so called UC, where each AP has to serve only a limited number of users. Moreover, we will present the method called successive lower bound maximization in order to maximizing the sum-rate and the energy efficiency. Because of the large number of antennas there is also the need of using hybrid architecture at each AP in order to reduce complexity and cost by using a small number of RF chains. With CF or UC, channel estimation and beamforming are locally evaluated, reducing the traffic load on the backhaul network. so, a comparison between a FD and hybrid architecture will be show. What we expect is that the FD architecture give better performances than the hybrid one. In the numerical results we will compare the performances in term of energy efficiency and sum-rate on Downlink and Uplink, with uniform and optimal power allocation and with a fully digital and hybrid architectures.

  This chapter is based on the journal article:

  – [5] M. Alonzo, S. Buzzi, A. Zappone, *et al.*, "Energy-efficient power control in cell-free and user-centric massive mimo at millimeter wave", *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 651–663, 2019

- **Chapter 4: Cell-Free and Distributed MIMO in Industry 4.0**
  In this chapter we will compare the performances of two different network architectures, CF (and UC) and D-MIMO for factory automation,

so this chapter focuses on industrial environments. D-MIMO is an architecture where the APs, connected to a CPU, are distributed in an area equipped with multiple antennas and in this case the beamformers and the channel estimates are not locally evaluated; this means that the transmitting antennas are distributed but the system behaves as if there were only one base station (BS), with more antennas. We will compare these two architectures in term of Signal-to-Interference-Plus-Noise-Ratio (SINR) and achievable rate by exploiting the FBLC [6]. This will be done in case of uniform and optimal power allocation, different beamforming schemes, such as maximum ratio transmission (MRT), full Zero-forcing (FZF), and partial zero-forcing (PZF) and different transmission modes, CFT, UCT, SAT, and JT.

This chapter is based on the journal article submitted to *IEEE Open Journal of the Communications Society*:

– M. Alonzo, P. Baracca, Saeed R. Khosravirad, and S. Buzzi, "Distributed and Cell-Free Massive MIMO Architectures for URLLC in Indoor Factory Environments"

## 1.3 Other Contributions

In addition to the paper listed above the author has other conferences publications. Below, only the abstract of those paper are reported.

### 1.3.1 Conferences

- [7] M. Alonzo, P. Baracca, S. R. Khosravirad, *et al.*, "URLLC for factory automation: An extensive throughput-reliability analysis of D-MIMO", in *WSA 2020; 24th International ITG Workshop on Smart Antennas*, 2020, pp. 1–6

Factory automation is one of the most challenging use cases for 5G-and-beyond mobile networks due to strict latency, availability and reliability constraints.

In this paper, communication with active actuators (ACs) inside an indoor industrial factory scenario is considered, and distributed multipleinput multiple-output (D-MIMO) systems are used in order to improve the communication reliability.We consider two transmission schemes, namely joint transmission, where all the distributed antennas are used for communicating with each AC, and single access point transmission, wherein, instead, each AC is served only by a subset of the active antennas. Link adaptation (LA) based on a proper choice of the modulation and coding scheme is also included in our analysis. Extensive numerical results are presented targeting various performance indicators, including the achievable sum spectral efficiency against the target block error rate, the number of ACs supported at a given reliability level, and finally, the effect of LA. The results demonstrate substantial performance improvements offered by D-MIMO in indoor factory scenarios.

- [8] M. Alonzo, S. Buzzi, and A. Zappone, "Energy-efficient downlink power control in mmWave cell-free and user-centric massive MIMO", in *2018 IEEE 5G World Forum*, IEEE, 2018

This paper considers cell-free and user-centric approaches for coverage improvement in wireless cellular systems operating at millimeter wave frequencies, and proposes downlink power control algorithms aimed at maximizing the global energy efficiency. To tackle the non-convexity of the problems, an interaction between sequential and alternating optimization is considered. The use of hybrid analog/digital beamformers is also taken into account. The numerical results show the benefits obtained from the power control algorithm, as well as that the user-centric approach generally outperforms the cell-free one.

- [9] M. Alonzo and S. Buzzi, "Cell-free and user-centric massive MIMO at millimeter wave frequencies", in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5. DOI: 10.1109/PIMRC.2017.8292302

In a cell-free (CF) massive MIMO architecture a very large number of distributed access points (APs) simultaneously and jointly serves a much smaller number of mobile stations (MSs); a variant of the cellfree technique is the user-centric (UC) approach, wherein each AP just decodes a reduced set of MSs, practically the ones that are received best. This paper introduces and analyzes the CF and UC architectures at millimeter wave (mmWave) frequencies. First of all, a multiuser clustered channel model is introduced in order to account for the correlation among the channels of nearby users; then, an uplink multiuser channel estimation scheme is described along with low-complexity hybrid analog/digital beamforming architectures. Interestingly, in the proposed scheme no channel estimation is needed at the MSs, and the beamforming schemes used at the MSs are channelindependent and have a very simple structure. Numerical results show that the considered architectures provide good performance, especially in lightly loaded systems, with the UC approach outperforming the CF one.

# Chapter 2

# Notion of Massive MIMO

This chapter refers to the book [10].
Today, almost all communications are wireless, this means that it is possible to access to the network almost everywhere thank to the cellular networks. As it has been said in the introduction, everything will be connected by exploiting a wireless connection and as a consequence the demand for wireless data connectivity will keep to increase in the future and this is the reason why there is the need of defining a new technology, in order to handle more data traffic. First of all, the concept of cellular network has to be clarified.

**Definition of cellular network**

Wireless communication is based on radio, so electromagnetic (EM) waves are used to carry information from a transmitter to one or multiple receivers. Since the EM waves propagate by following different paths from the transmitter, the signal energy spreads out and the signal reaches the receiver with less energy, since it decreases as the distance increases. To deliver information with EM waves with high received signal energy over a coverage area, a cellular network topology is needed [11]. A coverage area is divided into cells and in each cell there is a BS needed to enable wireless communication between a device and the network (see Fig.2.1). So, a cellular network is an architecture where there is a set of BSs distributed among cells and a set of UEs or MSs. The downlink (DL) refers to signal sent from a BS to served MSs, while Uplink (UL) is the transmission from MSs to their serving BS. A metric

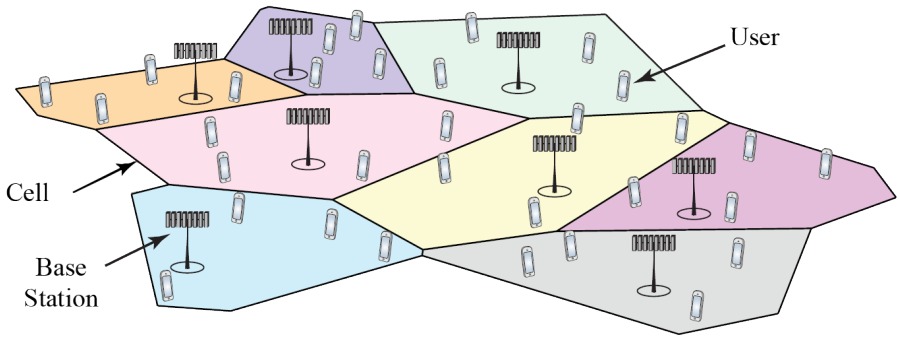



FIGURE 2.1: Traditional cellular network, where each BS covers a distinct geographical area and provides service to all users in it. The area is called a "cell" and is each cell is represented by a different color.

to asses the performances of conteporary and future cellular networks is the *area throughput*, measured in bit/s/km$^2$ and defined as:

$$\text{Area Throughput } [bit/s/km^2] = B \cdot D \cdot SE \tag{2.1}$$

where B [Hz] is the bandwidth, D [cell/km$^2$] is the average cell density and SE [bit/s/Hz/cell] is the spectral efficiency per cell. As we can see from the previous equation, there are three main components determining the area throughput and its enhancements; this means that, in order to increase the area throughput, it is possible to:

- use of larger bandwidth;
- employ more BSs;
- increase the SE per cell.

A way to increase the SE per cell is given by the use of Massive MIMO networks.

### 2.0.1 Definition of Massive MIMO

In this subsection we will provide a definition of the concept of massive MIMO, with the description of channel model, the DL and UL system model, channel hardening and SE with its bounds. As already discussed in this chapter, highly spectral efficient coverage tier[1] can be characterized as follows:

- exploit SDMA to achieve multiplexing gain by serving more MSs on the same time-frequency resources;
- more BS antennas than MSs per cell to enhance interference suppression. This implies that the number of BS antennas proportionally increases as the number of MSs increase.
- It operates in Time Division Duplex (TDD) mode to limit the channel state information (CSI) acquisition overhead.

**Definition of massive MIMO:** a massive MIMO network is a multicarrier network with C cells that operate according to TDD protocol; in this case DL and UL transmissions of different cells are synchronized. The *j*-th BS is equipped with $M \gg 1$ antennas, to achieve channel hardening, and communicates with K single-antenna users simultaneously. Thus, the antenna-MS ratio is M/K>1. Each BS locally process signal by using combining and precoding. An important aspect for massive MIMO networks is given by the fact that the frequency resources are divided into flat-fading subcarriers. Therefore, the *coherence bandwidth* $B_c$, which describes the frequency interval over which the channel responses, are almost constant, becomes really important too. One or multiple subcarriers are embedded in the coherence bandwidth, thus, the channel on andjacent subcarriers are almost equal. This guarantees that there is no the need to estimate channels for each subcarrier.

[1]Outdoor BSs that provide wide-area coverage, mobility support, and are shared between many MSs;

### 2.0.2 Channel model

The channel response between the $k$-th MS in the $c$-th cell and the $j$-th BS is denoted by $\mathbf{h}_{k,j,c} \in \mathbb{C}^M$, where each entry of this vector correspond to the channel between the MS to one of the BS's antenna. Since it has been assumed to operate with TDD, the channel on DL and UL are the same, for a coherence block[2]. UL channel is denoted by $\mathbf{h}_{k,j,c}$, whereas, DL channel is denoted by $(\mathbf{h}_{k,j,c})^H$. Since the channel response is a vector, it has a norm and direction which are both random variables in a fading channel. The channel model characterizes the distribution and statistical independence and dependence. A channel $\mathbf{h} \in \mathbb{C}^M$ is spatially uncorrelated if the channel gain $\|\mathbf{h}\|^2$ and the channel direction, $\mathbf{h}/\|\mathbf{h}\|$ are independent, where the channel direction is uniformly distributed over $\mathbb{C}^M$. If this not happens, the channel is spatially correlated. In this chapter, we will focus our attention to correlated Rayleigh fading channel, distributed as:

$$\mathbf{h}_{k,j,c} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}_M, \mathbf{R}_{k,j,c}\right) \tag{2.2}$$

where $\mathbf{R}_{k,j,c} \in \mathbb{C}^{M \times M}$ is the spatial correlation matrix, and it is a positive semi-definite matrix. In case of uncorrelated Rayleigh fading $\mathbf{R}_{k,j,c} = \beta_{k,j,c}\mathbf{I}_M$, where $\beta_{k,j,c}$ represent the large-scale fading for the $k$-th user, $j$-th BS in the $c$-th cell.

### 2.0.3 System model

Since, we have define the channel model, now it is possible to give the definition of system model for UL and DL.

#### Uplink

The Uplink transmission in a Massive MIMO system is illustrated in fig.2.2.

[2]Number of subcarriers and time samples over which the channel response can be approximated as constant and flat-fading

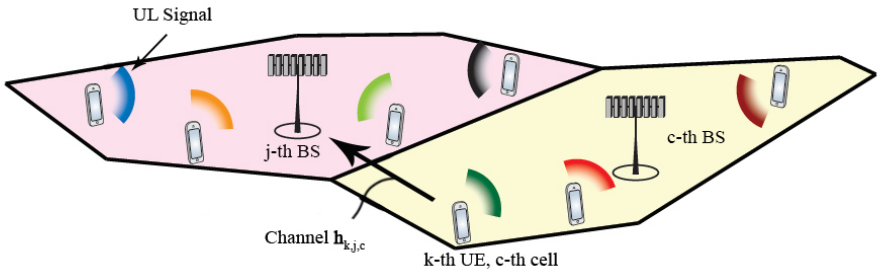



FIGURE 2.2: Uplink Massive MIMO transmission in the c-th cell and j-th cell. The channel vector between the k-th UE in c-th cell and j-th BS is denoted by $\mathbf{h}_{k,j,c}$.

The received signal, denoted by $\mathbf{y}_j \in \mathbb{C}$, at the $j$-th BS can be written as:

$$\begin{aligned}\mathbf{y}_j &= \sum_{c=1}^{C}\sum_{k=1}^{K_c}\mathbf{h}_{k,j,c}s_{k,c}^{UL} + \mathbf{n}_j = \\ &= \underbrace{\sum_{k=1}^{K_c}\mathbf{h}_{k,j,j}s_{k,c}^{UL}}_{\text{Useful signal}} + \underbrace{\sum_{c=1,c\neq j}^{C}\sum_{i=1}^{K_c}\mathbf{h}_{i,j,c}s_{i,c}^{UL}}_{\text{Inter-cell Interference}} + \underbrace{\mathbf{n}_j}_{\text{Noise}}\end{aligned} \tag{2.3}$$

where $\mathbf{n}_j \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_M, \sigma_{UL}^2\mathbf{I}_M)$ is the independent additive noise at the receiver. The signal to be sent from the $k$-th user in the $c$-th cell is denoted by $s_{k,c}^{UL} \in \mathbb{C}$, with power $p_{k,c} = \mathbb{E}\{|s_{k,c}^{UL}|^2\}$; the channel is constant for the same coherence block, whereas signals and noise are generated for each new sample. During the UL data transmission, the $j$-th BS has to apply a combining vector, $\mathbf{g}_{k,j} \in \mathbb{C}^M$ in order to separate the useful signal for $k$-th MS and its interference term, so we have:

$$\mathbf{g}_{k,j}^H\mathbf{y}_j = \underbrace{\mathbf{g}_{k,j}^H\mathbf{h}_{k,j,c}s_{k,j}^{UL}}_{\text{Useful Signal}} + \underbrace{\sum_{i=1,i\neq k}^{K_c}\mathbf{g}_{k,j}^H\mathbf{h}_{i,j,j}s_{i,j}^{UL}}_{\text{Intra-cell signals}} + \underbrace{\sum_{c=1,c\neq j}^{C}\sum_{i=1}^{K_c}\mathbf{g}_{k,j}^H\mathbf{h}_{i,j,c}s_{i,c}^{UL}}_{\text{Inter-cell Interference}} + \underbrace{\mathbf{g}_{k,j}^H\mathbf{n}_j}_{\text{Noise}} \tag{2.4}$$

### Downlink

The Downlink transmission in a Massive MIMO system is illustrated in fig.2.2. During DL phase, the $j$-th BS in $c$-th cell, transmit the following signal:

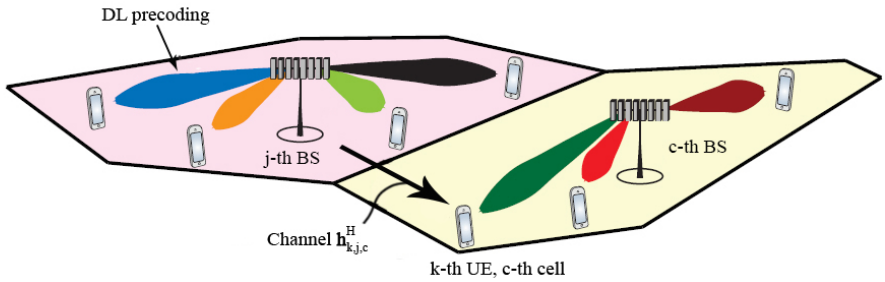



FIGURE 2.3: Downlink Massive MIMO transmission in the c-th cell and j-th cell. The channel vector between the k-th UE in c-th cell and j-th BS is denoted by $\mathbf{h}_{k,j,c}^{H}$.

$$\mathbf{x}_c = \sum_{k=1}^{K_c} \mathbf{q}_{k,c} s_{k,c}^{DL} \tag{2.5}$$

where $s_{k,c}^{DL} \sim \mathcal{N}(0, p_{k,c})$ is the downlink data signal to be sent to the $k$-th user in the $c$-th cell with power $p_{k,c}$. The spatial directivity of that signal is given by the precoding vector $\mathbf{q}_{k,c} \in \mathbb{C}^M$; the precoding vector has $\mathbb{E}\{\|\mathbf{q}_{k,c}\|^2\} = 1$, so, in this way $\mathbb{E}\{\|\mathbf{q}_{k,c} s_{k,c}^{DL}\|\} = p_{k,c}$, that is the transmit power allocated for $k$-th user. The received signal, denoted by $r_{k,c} \in \mathbb{C}$ at the $k$-th MS in the $j$-th cell is modelled as:

$$\begin{aligned}
\mathbf{r}_{k,j} &= \sum_{c=1}^{C} \mathbf{h}_{k,j,c}^{H} \mathbf{x}_j + n_{k,j} = \\
&= \sum_{c=1}^{C} \sum_{i=1}^{K_c} \mathbf{h}_{k,j,c}^{H} \mathbf{q}_{k,c} s_{k,c}^{DL} + n_{k,c} = \\
&= \underbrace{\mathbf{h}_{k,j,j}^{H} \mathbf{q}_{k,j} s_{k,j}}_{\text{Useful signal}} + \underbrace{\sum_{i=1}^{K_j} \mathbf{h}_{k,j,j}^{H} \mathbf{q}_{i,j} s_{i,j}}_{\text{Intra-cell interference}} + \underbrace{\sum_{l=1}^{C} \sum_{i=1, i \neq j}^{K_j} \mathbf{h}_{k,l,j}^{H} \mathbf{q}_{i,l} s_{i,l}}_{\text{Inter-cell interference}} + \underbrace{n_{k,j}}_{\text{Noise}}
\end{aligned} \tag{2.6}$$

where $n_{k,j} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_{DL}^2)$ is independent additive noise at the receiver. The channel is constant for the same coherence block, whereas signals and noise are generated for each new sample.

### 2.0.4 Channel hardening and favorable propagation

There two important aspects of multiantenna channel: *channel hardening* and *favorable propagation*. In this chapter we will provide only the definitions. With channel hardening it is possible to treat a fading channel as a determinstc one. This is useful to combat the small-scale fading and improve the DL channel gain estimation.

**Definition of channel hardening:** the channel response $\mathbf{h}_{k,j,j}$ provides (asymptotic) channel hardening if

$$\frac{\|\mathbf{h}_{k,j,j}\|^2}{\mathbb{E}\{\|\mathbf{h}_{k,j,j}\|^2\}} \to 1 \tag{2.7}$$

as $M \to \infty$. It is easy to see that the gain $\|\mathbf{h}_{k,j,j}\|^2$ of a channel response $\mathbf{h}_{k,j,j}$ is close to its mean when the number of antenna at each BS increases. In practice, is not relevant the asymptotic result, but how close to channel hardening we are when a lot of antennas are employed. Exploiting the concept of favorable propagation it is possible to make the direction of two user channels asymptotic orthogonal. This is useful when the interference between two user has to be mitigated and the SE is improved.

**Definition of favorable propagation**: asymptotic favorable propagation between two user channel, $\mathbf{h}_{i,j,c}$ and $\mathbf{h}_{k,j,j}$, to the $j$-th BS if

$$\frac{\mathbf{h}_{i,j,c}^H \mathbf{h}_{k,j,j}}{\sqrt{\mathbb{E}\{\|\mathbf{h}_{i,j,c}\|^2\}\mathbb{E}\{\|\mathbf{h}_{k,j,j}\|^2\}}} \to 0 \tag{2.8}$$

as $M \to \infty$. It is easy to see that the inner product of $\frac{\mathbf{h}_{i,j,c}}{\sqrt{\mathbb{E}\{\|\mathbf{h}_{i,j,c}\|^2\}}}$ and $\frac{\mathbf{h}_{k,j,j}}{\sqrt{\mathbb{E}\{\|\mathbf{h}_{k,j,j}\|^2\}}}$ goes asymptotically to zero. Since the channel norms grow as M grows, favorable propagation guarantees that the direction becomes orthogonal, but not

the channel responses of those users.

## 2.1 Channel estimation

In this section we describe the uplink training phase and provide a scheme of channel estimation, the minimum mean-squared error (MMSE).

### 2.1.1 Uplink training

In order to use efficiently a massive MIMO system, each BS has to estimate the channel response coming from the served user in the current coherence block. It is also important for the $j$-th BS to know the channel estimates of the other channels coming from users in cell j. This knowledge can be useful in order to mitigate the interference during the data transmission. Let $\tau_p$ the sampler used for uplink training phase in each coherence block. Each MS has to transmit a pilot sequence with length $\tau_p$; so, for the $k$-th user in the $j$-th cell, the pilot sequence can be denoted by $\phi_{k,j} \in \mathbb{C}^{\tau_p}$. This sequences have unit-magnitude element, so that $\|\phi_{k,j}\|^2 = \phi_{k,j}^H \phi_{k,j} = \tau_p$. Each element is rescaled by the UL power as $\sqrt{p_{k,j}}$ and then transmitted as the signal $s_{k,j}^{DL}$ in eq.(2.6). The received signal at the $j$-th BS, $\mathbf{Y}_j^p \in \mathbb{C}^{M \times \tau_p}$, can be written as:

$$\mathbf{Y}_j^p = \sum_{k=1}^{K_j} \sqrt{p_{k,j}} \mathbf{h}_{k,j,j} \phi_{k,j}^T + \sum_{c=1, c \neq j}^{C} \sum_{k=1}^{K_l} \sqrt{p_{i,c}} \mathbf{h}_{i,j,c} \phi_{i,c}^T + \mathbf{N}_j^p \tag{2.9}$$

where $\mathbf{N}_j^p \in \mathbb{C}^{M \times \tau_p}$ is the additive noise at the receiver with i.i.d. elements distributed as $\mathcal{N}_{\mathbb{C}}(0, \sigma_{UL}^2)$. $\mathbf{Y}_j^p$ is the observation available at the $j$-th BS and can be used to estimate the channel responses. To estimate the channel each BS needs to know which pilot sequence has been assigned to each user. Usually this assignment has to be done when a MS connect to the BS, by using, for example, a random process. Assuming that the $j$-th BS wants to estimate the channel response of the $k$-th user in the $c$-th cell. The BS has to multiply the

received signal $\mathbf{Y}_j^p$ with the pilot sequence $\phi_{k,l}$, so we have:

$$\mathbf{y}_{k,j,c}^p = \mathbf{Y}_j^p \phi_{i,c}^* = \sum_{c'=1}^{L} \sum_{i'=1}^{K_{c'}} \sqrt{p_{i',c'}} \mathbf{h}_{i',j,c'} \phi_{i',c'}^T \phi_{i,c}^* + \mathbf{N}_j^p \phi_{i,c}^* \tag{2.10}$$

For the BS's own cell it is possible to rewrite the eq.(2.10) as follows:

$$\begin{aligned} \mathbf{y}_{k,j,j}^p &= \mathbf{Y}_j^p \phi_{k,j}^* = \\ &= \underbrace{\sqrt{p_{k,j}} \mathbf{h}_{k,j,j} \phi_{k,j}^T \phi_{k,j}^*}_{\text{Useful pilot}} + \underbrace{\sum_{i=1, i\neq k}^{K_j} \sqrt{p_{i,j}} \mathbf{h}_{i,j,j} \phi_{i,j}^T \phi_{k,j}^*}_{\text{Intra-cell Pilots}} + \\ &+ \underbrace{\sum_{c=1, c\neq j}^{K} \sum_{i=1, i\neq k}^{K_j} \sqrt{p_{i,c}} \mathbf{h}_{i,j,c} \phi_{i,c}^T \phi_{k,j}^*}_{\text{Inter-cell Pilots}} + \underbrace{\mathbf{N}_j^p \phi_{i,c}^*}_{\text{Noise}} \end{aligned} \tag{2.11}$$

The Intra-cell and inter-cell interference in eq.(2.11) contains the inner product between the pilot of $k$-th and another user in the $l$-th cell. If that inner product is zero, than the interference vanishes, and this happens iff the pilot sequence are of two users are orthogonal (i.e. $\phi_{k,c}^T \phi_{k,j}^* = 0$). We would assign orthogonal pilots, for all users, but due to the finite length of coherence blocks, which imposes the constraint $\tau_p \leq \tau_c$, this is not possible in practice. In this chapter, we assume to use a set of $\tau_p$ mutually orthogonal pilots; this pilots can be seen as columns of a *pilot book* $\mathbf{\Phi} \in \mathbb{C}^{\tau_p \times \tau_p}$, which has to satisfy $\mathbf{\Phi}^H \mathbf{\Phi} = \tau_p \mathbf{I}_p$[3]. Now, we can define the set

$$\mathcal{P}_{k,j} = \left\{ (i,c) : \ \phi_{i,c} = \phi_{k,j}, \ c = 1, \ldots, C, \ i = 1, \ldots, K_c \right\} \tag{2.12}$$

with the indices of MSs that use the same pilots as the $k$-th MS in the $j$-th cell. Thus, $(l,i) \in \mathcal{P}_{k,j}$ implies that the $i$-th MS in the $l$-th cell uses the same pilot as

[3]It is better to set $\tau_p \geq \max_l K_l$ so that each BS can use different pilot for its served MSs. In this way the strongest interference coming from users in the own cell.

$k$-th MS in the $j$-th cell. Now, eq.(2.11) can be rewritten as:

$$\mathbf{y}_{k,j,j}^{p} = \underbrace{\sqrt{p_{k,j}}\tau_p \mathbf{h}_{k,j,j}}_{\text{Useful pilot}} + \underbrace{\sum_{(i,c)\in\mathcal{P}_{k,j}\setminus(k,j)} \sqrt{p_{i,c}}\tau_p \mathbf{h}_{i,j,c}}_{\text{Interfering Pilots}} + \underbrace{\mathbf{N}_j^p \phi_{i,c}^*}_{\text{Noise}} \tag{2.13}$$

Note that $\mathbf{y}_{k,j,j}^{p} = \mathbf{y}_{i,j,c}^{p}$, $\forall (c,i) \in \mathcal{P}_{k,j}$, since these MSs are using the same pilot; moreover, $\mathbf{N}_j^p \phi_{k,j}^* \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}_M, \sigma_{UL}^2 \tau_p \mathbf{I}_M)$, since the pilot are deterministic and $\|\phi_{k,j}\| = \tau_p$. In literature, there are different channel estimation schemes, but in this thesis we will only consider the Minimum mean square error (MMSE). Let us focus our attention to the case of spatially uncorrelated channel, i.e., $\mathcal{CN}(\mathbf{0}_M, \beta_{k,j,c}\mathbf{I}_M)$, the MMSE estimates can be easily written as:

$$\hat{\mathbf{h}}_{k,j,c} = \frac{\sqrt{p_{k,c}}\beta_{k,j,c}}{\sum\limits_{(i,c')\in\mathcal{P}_{k,c}} p_{i,c'}\beta_{i,j,c'} + \sigma_{UL}^2} \mathbf{y}_{k,j,c}^{p} = \alpha_{k,j,c}\mathbf{y}_{k,j,c}^{p} \tag{2.14}$$

## 2.2 Spectral efficiency: DL bound for Massive MIMO

In this section we will provide the definitions of upper and lower bounds for the SE on the DL for a massive MIMO network. This results has been useful for the analysis addressed in Chapter 4. Each BS has to transmit payload data to its users during the DL phase by using a linear precoding. Precoding vector has to construct such that $\mathbb{E}\{\|\mathbf{q}_{k,j}\|\} = 1$, in this way $p_{k,j}$ is the signal power allocated to the $k$-th user. Recall that $s_{k,j}^{DL}$ represents the information to be sent to the $k$-th user.

### 2.2.1 DL upper bound

The upper bound can be easily computed. If we consider eq.(2.6), the upper bound can be written as:

$$\mathrm{SE}_{k,j}^{DL} = \frac{\tau_d}{\tau_c}\mathbb{E}\left[\log_2\left(1 + \mathrm{SINR}_{k,j}^{DL}\right)\right] \tag{2.15}$$

where the quantity $\frac{\tau_d}{\tau_c}$ is the fraction of samples for coherence block used for DL data, where $\tau_d = \tau_c - \tau_p - \tau_u$, wherein $\tau_u$ is the number of samples used for UL data, while $\text{SINR}_{k,j}^{DL}$ is defined as:

$$\text{SINR}_{k,j}^{DL} = \frac{\eta_{k,j}|\mathbf{h}_{k,j,j}^H \mathbf{q}_{k,j}|^2}{\sum\limits_{l=1}^{C} \sum\limits_{i=1,(i,l)\neq(k,j)}^{K_l} \eta_{i,l}|\mathbf{h}_{k,l,j}^H \mathbf{q}_{i,l}|^2 + \sigma_{DL}^2} \tag{2.16}$$

where $\eta_{k,j}$ is the power allocated for the $k$-th user by the $j$-th BS.

### 2.2.2 DL lower bound

The desired signal intended to the $k$-th user in the $j$-th cell can be written as $g_{k,j} = \mathbf{h}_{k,j,j}^H \mathbf{q}_{k,j}$; the user does not know a priori this quantity but it can be approximated by using its mean value. The received DL signal, $y_{k,j}$, at the $k$-th user can be written as:

$$\begin{aligned} y_{k,j} = & \underbrace{\mathbb{E}\{\mathbf{h}_{k,j,j}^H \mathbf{q}_{k,j}\} s_{k,j}^{DL}}_{\text{Useful signal over average channel}} + \underbrace{\left(\mathbf{h}_{k,j,j}^H \mathbf{q}_{k,j} - \mathbb{E}\{\mathbf{h}_{k,j,j}^H \mathbf{q}_{k,j}\}\right) s_{k,j}^{DL}}_{\text{Useful signal over unknown channel}} + \\ & + \underbrace{\sum_{i=1,i\neq k}^{K_j} \mathbf{h}_{k,j,j}^H \mathbf{q}_{i,j}}_{\text{Intra-cell interference}} + \underbrace{\sum_{l=1,l\neq j}^{C} \sum_{i=1}^{K_l} \mathbf{h}_{k,l,j}^H \mathbf{w}_{i,l} s_{i,l}^{DL}}_{\text{Inter-cell interference}} + \underbrace{n_{k,j}}_{\text{Noise}} \end{aligned} \tag{2.17}$$

where the first term is the useful signal over the deterministic average precoded channel, while the other terms are unknown to the user since they are random variables. An achievable SE can be evaluated by treating these terms as noise in signal detection. Then it is possible to get a capacity bound, called *hardening bound*. Of course, this bound is valid for any beamformer and channel estimation scheme. The DL bound for the $k$-th user in the $j$-th cell is written as $\underline{\text{SE}}_{k,j}^{DL} = \frac{\tau_d}{\tau_c} \log_2(1 + \underline{\text{SINR}}_{k,j}^{DL})$ [bit/s/Hz] with

$$\underline{\text{SINR}}_{k,j}^{DL} = \frac{\eta_{k,j}|\mathbb{E}\{\mathbf{q}_{k,j}^H \mathbf{h}_{k,j,j}\}|^2}{\sum\limits_{l=1}^{C} \sum\limits_{i=1}^{K_l} \eta_{i,l}\mathbb{E}\{|\mathbf{q}_{i,l}^H \mathbf{h}_{k,l,j}|^2\} - \eta_{k,j}|\mathbb{E}\{\mathbf{q}_{k,j}^H \mathbf{h}_{k,j,j}\}|^2 + \sigma_{DL}^2} \tag{2.18}$$

where the numerator is the gain of useful signal receiver over average precoded channel. Regarding the denominator, the first term is the received power of all signals, the second one, instead, subtracts the part of the useful signal that appear at the numerator, and the third and last term is the noise variance. The SE bound can be evaluated numerically, for each beamformer and channel estimation scheme, by computing all the averages with Montecarlo method. In case of MRT beamformer, $\mathbf{q}_{k,c} = \hat{\mathbf{h}}_{k,j,c}/\|\hat{\mathbf{h}}_{k,j,c}\|^2$, in case of spatially uncorrelated channel, i.e., $\mathbf{h}_{k,j,c} \sim \mathcal{CN}(\mathbf{0}_M, \beta_{k,j,c}\mathbf{I}_M)$, and MMSE channel estimation it is possible to write that lower bound in a closed form as follows:

$$\mathrm{SINR}_{k,j}^{DL} = \frac{p_{k,j}\eta_{k,j}\beta_{k,j,j}\alpha_{k,j,j}M^2}{\sum\limits_{c=1}^{C}\sum\limits_{i=1}^{K_c}\eta_{k,l}\sqrt{p_{k,l}}\beta_{k,j,l}\alpha_{i,l,l}\beta_{i,l,l}M + \\ + \sum\limits_{(i,l)\in\mathcal{P}_{k,j}\setminus(k,j)}\eta_{k,l}\sqrt{p_{k,j}}\beta_{k,j,l}\alpha_{i,l,l}\beta_{i,l,l}M^2 + \sigma_{DL}^2} \tag{2.19}$$

# Chapter 3

# Cell-Free Massive MIMO at mmWave frequncies

Referred articles published in the *IEEE 5G World Forum (5GWF). IEEE, 2018 and in the IEEE Transactions on Green Communications and Networking 3.3, 2019*.

## 3.1 Introduction

Future 5G wireless systems will heavily rely on the use of large-scale antenna arrays, a.k.a. massive MIMO, and of carrier frequencies above 10 GHz, the so called mmWave frequencies [4]. Indeed, on one hand, the use of massive MIMO permits serving several users on the same time-frequency resource, while, on the other one, mmWave carrier frequencies will enable the use of much larger bandwidths, at least on short-distances (up to about one hundred meters). The combined use of massive MIMO systems along with mmWave frequencies is indeed one of the key technological enablers of the envisioned wireless Gbit/s experience [12]. For conventional sub-6 GHz frequencies, a new communications architecture, named CF massive MIMO, has been recently introduced in [13], [14], in order to alleviate the cell-edge problem and thus increase the system performance of unlucky users that happen to be located very far from their serving AP. In the CF architecture, instead of having few base stations with massive antenna arrays, a very large number

of simple APs randomly and densely deployed serve a much smaller number of MSs. This approach has some similarities with the CoMP [15], [16] and network MIMO [17]; these technologies exploit coordinate beamforming, scheduling, and joint transmission using multiple distributed antennas and/or base stations, with the aim of mitigating the interference and achieving diversity gains. There are however some key differences between these strategies and the CF massive MIMO approach discussed in this chapter. In particular, we have that: (a) the CF massive MIMO strategy fully leverages the advantages and unique features of massive MIMO, such as the channel hardening effect and the use of the TDD protocol;and (b) the CF massive MIMO also makes a more limited use of the backhaul link since the channel coefficients are locally estimated at the APs and are mutually shared. Moreover, beamformers are computed at the APs using locally available information. The CF massive MIMO architecture can be thus thought as the scalable way of implementing distributed network MIMO deployments. In the CF architecture described in [13], [14], single-antenna APs and MSs are considered, all the APs serve all the MSs, all the APs are connected through a backhaul link to a CPU. Nevertheless every AP performs channel estimation locally, and channel estimates are not sent to the CPU, but are locally exploited. In particular, for the downlink communication phase, the CPU sends to the APs the data symbols to be sent to all the MSs. Then, for the uplink communication phase, the APs use the backhaul to send the sufficient statistics for all the MSs to the CPU, which combines them and performs uplink data decoding. In [18] a UC variant of the CF approach is introduced, wherein each APs, instead of serving all the MSs in the considered area, just serves the ones that he receives best; the results of [18] show that the UC approach provides savings on the required backhaul capacity and, also, provides better data-rates to the vast majority of the users. The paper [19] analyzes a CF massive MIMO system from the point of view of its energy consumption, taking into account the power consumed by the backhaul links, the number of active APs and the number of antennas at each AP. The paper also embraces the UC philosophy by considering AP-MS selection schemes aimed at maximizing the

system energy efficiency. While previous works deal with the case in which the APs and the MSs are both equipped with one antenna (with the exception of [19] which considers multiple antennas at the APs), in [20] the CF and UC architectures are generalized to the case in which both the APs and the MSs are equipped with multiple antennas. This generalization is not trivial and indeed the system proposed in [20], despite the use of multiple antennas at the MSs and the use of a multiplexing order larger than one, develops a beamforming scheme at the APs that does not require channel estimation at the MSs, which adopt a channel-independent beamformer. While [20] considers a simple uniform power allocation, in [21], [22] power control procedures aimed at rate maximization are proposed. The paper [23] investigates the use of pilot signals on the downlink, so as to enable channel estimation at the MSs; even though such an approach provides some performance improvement, it contradicts the TDD-based "Massive MIMO philosophy" wherein no channel estimation is required at the MSs. A compute-and-forward approach to CF massive MIMO was then proposed in [24] in order to reduce the load on the backhaul links. All of the above cited papers consider the case in which a conventional sub-6 GHz carrier frequency is used. On the other hand, as already stated, mmWave frequencies will play a big role in future wireless cellular systems due to availability of wide amounts of unused bandwidths [25]. The design of a wireless cellular system operating at mmWave poses different and new challenges with respect to conventional sub-6 GHz frequencies, due to different propagation mechanisms [25], to the need of using large antenna arrays at both side of the links to counteract the increased path-loss [26], and to the presence of hardware constraints that prevent the realization of FD precoding and postcoding beamforming structures [27]. Despite these tough challenges, the attractiveness of the mmWave frequencies for cellular communications has led to intense research efforts in the last few years. Due to the difficulty of realizing FD beamforming architectures at mmWave with antenna arrays of large size, the CF massive MIMO architecture at these high frequencies appears particularly attractive, mainly for two reasons; first of all, it substitutes large co-located antennas with several APs with antenna arrays

of smaller dimension, and, thus, with less hardware complexity; second, due to the limited range of mmWave communications, the distributed AP dense deployment alleviates the cell-edge problem and creates path-diversity, which is precious since signal blockages may frequently happen at these frequencies. Energy efficiency is another key topic that should be considered when designing modern wireless communication systems. Indeed, increasing the bit-per-Joule energy efficiency is regarded as a key requirement of future 5G networks [28]. A recent survey on the most promising energy-efficient techniques for 5G has recently appeared in [29]; in that paper, radio resource allocation is identified, among other techniques, as a relevant approach to improve the energy efficiency of future wireless networks. Recent contributions on energy-efficient radio resource allocation for 5G are [30]–[33]. Motivated by the above discussion, this chapter considers details a CF and UC massive MIMO wireless system operating at mmWave and focuses on resource allocation strategies maximizing the global energy efficiency. To the best of authors' knowledge, this chapter is based on the first paper that considers the CF and UC massive-MIMO deployments at mmWave; preliminary results on this issue have been reported in the conference papers [8] and [9], wherein it has also been shown that the UC approach generally outperforms the CF approach. The contribution of this chapter can be summarized as follows. First of all, we introduce a multiuser mmWave channel model that permits taking into account channel correlation for close users. Building upon the clustered channel model [34] widely used at mmWave frequencies, we extend this model to take into account the fact that if several APs and MSs are in the same area, their channels must be built using the same set of scatterers; adopting this model, users that are very close will receive beams with very close direction of arrival and so channel correlation for nearby users is intrinsically taken into account. Then, we study the UC and CF approaches at mmWave frequencies; we assume that both the APs and MSs are equipped with multiple antennas, use HY analog-digital partial Zero-Forcing (ZF) beamforming at the APs, while a very simple 0-1 beamforming architecture, independent of

the channel estimate, is used at the MSs. We present simulation results for scenarios that can be representative of a lightly-loaded system, with $M = 80$ APs and $K = 6$ MSs transmitting on the same time-frequency slot, and of a highly-loaded system, with $M = 80$ APs and $K = 16$ MSs. Results show that over a bandwidth of 200 MHz and with a maximum transmit power of 0.1 W at each AP, taking into account channel estimation errors and using low-complexity beamforming structures, the downlink average rate per user is 1.5 Gbit/s in the lightly-loaded scenario and 400 Mbit/s in the heavily-loaded situation. Likewise, in the lightly-loaded scenario, the average uplink rate-per-user is about 1 Gbit/s. This chapter is organized as follows. Next section is devoted to the discussion of the system and to the description of the processing (it involves uplink training, downlink data transmission and uplink data transmission); Section 3.6 concerns global energy efficiency maximization; Section 3.7 concerns the sum-rate optimization under (theoretical) perfect ZF beamforming; Section 3.8 contains details about the used multiuser channel model and the discussion of the numerical results. Finally, concluding remarks are given in Chapter 5.

## 3.2 System model and communication protocol

We consider an area where $K$ MSs and $J$ APs are randomly located. The APs are connected by means of a backhaul network to a CPU wherein data-decoding is performed (see fig. 3.1). Communications take place on the same frequency band; downlink and uplink are separated through TDD[4]. The communication protocol is made of three different phases: uplink training, downlink data transmission and uplink data transmission. During the uplink training phase, the MSs send pilot sequences to the APs and each AP estimates the channels; during the second phase the APs use the channel estimates to perform pre-coding and transmit the data symbols; finally, in the third phase the MSs send uplink data symbols to the APs. While in the CF approach all

[4]In TDD, using calibrated hardware, the uplink channel is the reciprocal of downlink channel.

the APs simultaneously serve all the MSs (a fully-cooperative scenario), in the UC approach each AP serves a pre-determined number of MSs, say $N$, and in particular the ones that it receives best. We assume that each AP (MS) is

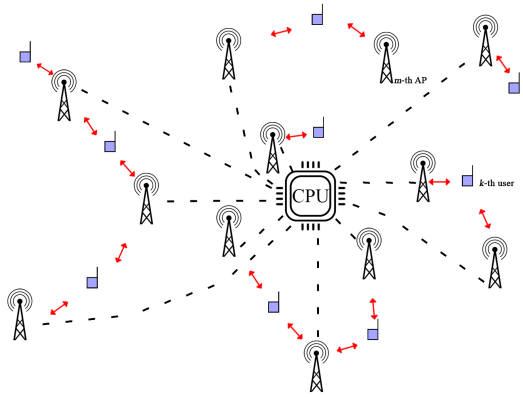



FIGURE 3.1: A CF massive MIMO network deployment.

equipped with an antenna array with $N_{AP}$ ($N_{MS}$) elements. The $(N_{AP} \times N_{MS})$-dimensional matrix $\mathbf{H}_{k,j}$ denotes the channel matrix between the $k$-th user and the $j$-th AP. Details about the channel model will be reported in Section 3.8

Additionally, we assume that each MS employs a very simple 0-1 beamforming structure; in particular, denoting by $P$ the multiplexing order, namely the number of parallel streams sent to a given receiver, the $(N_{MS} \times P)$-dimensional beamformer used at the $k$-th MS receiver is denoted by $\mathbf{L}_k$ and is defined as $\mathbf{L}_k = \mathbf{I}_P \otimes \mathbf{1}_{N_{MS}/P}$, denoting with $\mathbf{1}_{N_{MS}/P}$ an all-1 vector of length $N_{MS}/P$. Otherwise stated, we assume that the MS receive antennas are divided in $P$ disjoint groups of $N_{MS}/P$ elements, and the received data collected at the antennas of each group are simply summed together. It is APs' task, based on the uplink channel estimates and exploiting the TDD channel reciprocity, to ensure that the summed samples are, at least approximately, aligned in phase. Similarly when considering uplink transmission, the antennas in each group send the same signal with the same phase. We describe now the three phases of the communication protocol.

## 3.3 Uplink training

During the uplink training the MSs transmit pilot sequences in order to enable channel estimation at the APs. Let $\tau_c$ be the length of the channel coherence time and $\tau_p$ be the length of uplink training phase, both in discrete time samples. Of course we must have $\tau_p < \tau_c$. We define by $\mathbf{\Phi}_k \in \mathcal{C}^{P \times \tau_p}$ the matrix containing on its rows the pilot sequences sent by the $k$-th MS. We assume that $\mathbf{\Phi}_k \mathbf{\Phi}_k^H = \mathbf{I}_P$, i.e. the rows of $\mathbf{\Phi}_k$ are orthogonal, but no orthogonality is required for the pilot sequences assigned to other MSs[5]. Obviously, using orthogonal pilots *tout court* would lead to a system immune to pilot contamination, but this would put a limit on the maximum value of the product $KP$ that could be accommodated in the channel coherence time. The received signal at the $j$-th AP in the $\tau_p$ signaling intervals devoted to uplink training can be cast in the following $N_{AP} \times \tau_p$-dimensional matrix $\mathbf{Y}_j$:

$$\mathbf{Y}_j = \sum_{k=1}^{K} \sqrt{p_k} \mathbf{S}_{k,j} \mathbf{\Phi}_k + \mathbf{W}_j, \tag{3.1}$$

where $\mathbf{S}_{k,j} = \mathbf{H}_{k,j} \mathbf{L}_k$, $\mathbf{W}_j$ is the matrix of thermal noise samples, whose entries are assumed to be i.i.d. $\mathcal{CN}(0, \sigma_w^2)$ RVs. In the following we briefly outline the structure of the Linear Minimum Mean Square Error (LMMSE) channel estimator. Defining $\mathbf{y}_j = \text{vec}(\mathbf{Y}_j)$, $\mathbf{w}_j = \text{vec}(\mathbf{W}_j)$, $\mathbf{s}_{k,j} = \text{vec}(\mathbf{S}_{k,j})$, we obtain the vectorized model:

$$\mathbf{y}_j = \sum_{k=1}^{K} \sqrt{p_k} \mathbf{A}_k \mathbf{s}_{k,j} + \mathbf{w}_j \tag{3.2}$$

[5]Of course, when $KP \leq \tau_p$ it would be possible to assign to all the MSs mutually orthogonal pilot sequences. In general, the pilot sequences contain symbols from the same constellation as the data symbols, e.g. QAM symbols or similar. In this chapter, however, for the sake of simplicity, we assume that the pilot sequences are binary random sequences, and we just require that each matrix $\mathbf{\Phi}_k$ has orthogonal rows.

**Algorithm 1** Block Coordinate Descent for Subspace Decomposition Algorithm for Hybrid Beamforming; the algorithm input is $\mathbf{Q}_j^{opt}, N_{AP}, N_{AP}^{RF}, K, P$.

1: Initialize $I_{max}$ and set iter=0;
2: Set arbitrary $\mathbf{Q}_j^{RF,0}$
3: **repeat**
4: Update $\mathbf{Q}_j^{BB} = (\mathbf{Q}_j^{RF,iter\,H}\mathbf{Q}_j^{RF,iter})^{-1}\mathbf{Q}_j^{RF,iter\,H}\mathbf{Q}_j^{opt}$
5: Set $\phi_{iter} = \text{angle}\Big[\mathbf{Q}_j^{opt}\mathbf{Q}_j^{BB,iter+1\,H}(\mathbf{Q}_j^{BB,iter+1}\mathbf{Q}_j^{BB,iter+1\,H})^{-1}\Big]$
6: Update $\mathbf{Q}_j^{RF,iter} = \frac{1}{\sqrt{N_{AP}}}e^{i\phi_{iter}}$
7: $iter = iter + 1$;
8: **until** convergence or $iter = I_{max}$

with $\mathbf{A}_k = \Phi_k^T \otimes \mathbf{I}$. Since we need to estimate a vector, we process $\mathbf{y}_j$ by a matrix $\mathbf{V}_{k,j}^H$, i.e. $\hat{\mathbf{s}}_{k,j} = \mathbf{V}_{k,j}^H\mathbf{y}_j$. Then, the MSE is:

$$\begin{aligned}
&\mathbb{E}[\| \mathbf{V}_{k,j}^H\mathbf{y}_j - \mathbf{s}_{k,j} \|^2] = \text{tr}(\mathbf{V}_{k,j}^H\mathbb{E}[\mathbf{y}_j\mathbf{y}_j^H]\mathbf{V}_{k,j}) + \\
&+ \mathbb{E}[\| \mathbf{s}_{k,j} \|^2] - \mathbb{E}[2\Re\{\text{tr}(\mathbf{s}_{k,j}^H\mathbf{V}_{k,j}^H\mathbf{y}_j)\}] = \\
&= \text{tr}\left(\mathbf{V}_{k,j}^H\left(\sum_{l=1}^{K} p_k\mathbf{A}_l\mathbf{A}_l^H + \sigma^2\mathbf{I}\right)\mathbf{V}_{k,j}\right) + \\
&+ \mathbb{E}[\| \mathbf{s}_{k,j} \|^2] - \sqrt{p_k}\text{tr}(\mathbf{V}_{k,j}^H\mathbf{A}_k + \mathbf{V}_{k,j}^T\mathbf{A}_k^*)
\end{aligned} \tag{3.3}$$

Recalling that ([35, Ch. 4]) $\nabla_{\mathbf{Z}^*}\text{tr}(\mathbf{Z}^T\mathbf{M}) = \mathbf{0}$, $\nabla_{\mathbf{Z}^*}\text{tr}(\mathbf{Z}^H\mathbf{M}) = \mathbf{M}$, and $\nabla_{\mathbf{Z}^*}\text{tr}(\mathbf{Z}^H\mathbf{M}\mathbf{Z}) = \mathbf{M}\mathbf{Z}$, setting the gradient of the MSE with respect to the complex matrix $\mathbf{V}_{k,j}^*$ equal to zero and solving for $\mathbf{V}_{k,j}$, we find the LMMSE estimator:

$$\mathbf{V}_{k,j}^{LMMSE} = \sqrt{p_k}\Big(\sum_{l=1}^{K} p_l\mathbf{A}_l\mathbf{A}_l^H + \sigma^2\mathbf{I}\Big)^{-1}\mathbf{A}_k\,. \tag{3.4}$$

## 3.4 Downlink data transmission

After the first phase, the generic $j$-th AP has an estimate of the quantities $\mathbf{S}_{k,j}$, for all $k = 1, \ldots, K$. In order to transmit data on the downlink, a ZF precoder is considered. The precoding matrix $\mathbf{Q}_{k,j}$ is designed as follows. First of all,

the $(N_{AP} \times KP)$-dimensional matrix $\mathcal{G}_j = [\mathbf{S}_{1,j} \dots \mathbf{S}_{K,j}]$ is built; then, we have

$$\mathbf{Q}_{k,j} = (\mathcal{G}_j^H \mathcal{G}_j)^{-1} \mathbf{S}_{k,j}. \tag{3.5}$$

Finally, each precoding matrix is normalized as follows:

$$\mathbf{Q}_{k,j} = \frac{\mathbf{Q}_{k,j}}{\sqrt{\mathrm{tr}(\mathbf{Q}_{k,j}\mathbf{Q}_{k,j}^H)}}\ , \tag{3.6}$$

$\forall k = 1, \dots, K, \forall j = 1, \dots, J$.

The previously described beamforming matrix requires a FD implementation, which presumes the use of a number of RF chains equal to the number of transmit antennas. It is well-known that at mmWave frequencies hardware complexity constraints usually prevent the use of FD architectures, and thus HY beamforming structures have been proposed, where a number of RF chains $N_{AP}^{RF} < N_{AP}$ is used. Here, we exploit the *Block Coordinate Descent algorithm* [36] in order to decompose our beamformer in the cascade of a FD (baseband) one, represented by a $(N_{AP}^{RF} \times P)$-dimensional matrix and of an analog one, represented by a $(N_{AP} \times N_{AP}^{RF})$-dimensional matrix whose entries have all constant norm. While the baseband beamformer is MS-dependent, the analog beamformer at the AP is unique and is used for transmitting to all the users. In particular, at the generic $j$-th AP, the following matrix, of dimension $N_{AP} \times KP$, is formed[6]

$$\mathbf{Q}_j^{opt} = \left[\mathbf{Q}_{1,j}, \dots, \mathbf{Q}_{K,j}\right]\ , \tag{3.7}$$

and used as an input to the algorithm 1 in order to obtain the $N_{AP} \times N_{AP}^{RF}$ matrix $\mathbf{Q}_j^{RF}$ and the $N_{AP}^{RF} \times KP$ matrix

$$\mathbf{Q}_j^{BB} = \left[\mathbf{Q}_{1,j}^{BB}, \dots, \mathbf{Q}_{K,j}^{BB}\right]\ ,$$

[6]In the UC approach, to be detailed in the following, the matrix $\mathbf{Q}_j^{opt}$ is formed using the beamformers relative to the MS actually served by the $j$-th AP.

containing the MS-specific baseband beamformers to be used at the $j$-th AP. At the generic $j$-th AP, we will thus have as many digital beamformers as the MSs to transmit to, and only one analog beamformer, that will be used to transmit jointly to all the users. Although we are here considering, for the sake of simplicity, only one single frequency, it should be noted that with a multicarrier modulation a single analog beamformer must be used for all the subcarriers, or, if complexity permits, for each properly defined subset of contiguous subcarriers [36].

### 3.4.1 The CF approach

In this case, all the APs serve all the MSs, so the transmitted signal from the $j$-th AP in the $n$-th sample interval is:

$$\mathbf{s}_j^{CF}(n) = \sum_{k=1}^{K} \sqrt{\eta_{j,k}} \mathbf{Q}_{k,j} \mathbf{x}_k^{DL}(n), \tag{3.8}$$

where $\mathbf{x}_k^{DL}(n)$ is the data symbol intended for the $k$-th MS, and $\eta_{j,k}$ is a scalar coefficient taking into account the power used by the $j$-th AP to transmit towards the $k$.th MS. The $k-th$ MS receives the following $(N_{MS} \times 1)$-dimensional vector:

$$\begin{aligned} \mathbf{y}_k^{CF}(n) &= \sum_{j=1}^{J} \mathbf{H}_{k,j}^H \mathbf{s}_j^{CF}(n) + \mathbf{z}_k(n) = \\ &= \sum_{j=1}^{J} \sqrt{\eta_{j,k}} \mathbf{H}_{k,j}^H \mathbf{Q}_{k,j} \mathbf{x}_k^{DL}(n) + \\ &+ \sum_{l=1, l\neq k}^{K} \sum_{j=1}^{J} \sqrt{\eta_{j,l}} \mathbf{H}_{k,j}^H \mathbf{Q}_{l,j} \mathbf{x}_l^{DL}(n) + \mathbf{z}_k(n), \end{aligned} \tag{3.9}$$

where $\mathbf{z}_k(n)$ is the additive thermal noise distributed as $\mathcal{CN}(0, \sigma_z^2)$. A soft estimate of the $k$-th MS data symbol is thus formed as

$$\widehat{\mathbf{x}}_k^{DL,CF}(n) = \mathbf{L}_k^H \mathbf{y}_k^{CF}(n). \tag{3.10}$$

### 3.4.2 The UC approach

In this case the APs are assumed to serve a pre-determined, fixed number of MSs, say $N$; in particular, we assume that the generic $j$-th AP serves the $N$ MSs whose channels have the largest Frobenious norms. We denote by $\mathcal{K}(j)$ the set of MSs served by the $j$-th AP. Given the sets $\mathcal{K}(j)$, for all $m = 1, \ldots, M$, we can define the set $\mathcal{J}(k)$ of the APs that communicate with the $k$-th user:

$$\mathcal{J}(k) = \{j : k \in \mathcal{K}(j)\}. \tag{3.11}$$

In this case the transmitted signal from the $j$-th AP is written as:

$$\mathbf{s}_j^{UC}(n) = \sum_{k \in \mathcal{K}(j)} \sqrt{\eta_{j,k}} \mathbf{Q}_{k,j} \mathbf{x}_k^{DL}(n). \tag{3.12}$$

The received signal at the $k$-th MS is expressed now as:

$$\begin{aligned} \mathbf{y}_k^{UC}(n) &= \sum_{j=1}^{J} \mathbf{H}_{k,j}^H \mathbf{s}_j^{UC}(n) + \mathbf{z}_k(n) = \\ &= \sum_{j \in \mathcal{J}(k)} \sqrt{\eta_{j,k}} \mathbf{H}_{k,j}^H \mathbf{Q}_{k,j} \mathbf{x}_k^{DL}(n) + \\ &+ \sum_{l=1, l \neq k}^{K} \sum_{j \in \mathcal{J}(l)} \sqrt{\eta_{j,l}} \mathbf{H}_{k,j}^H \mathbf{Q}_{l,j} \mathbf{x}_l^{DL}(n) + \mathbf{z}_k(n), \end{aligned} \tag{3.13}$$

As before the $N_{MS}$-dimensional vector $\mathbf{z}_k(n)$ represents the thermal noise at the $k$-th MS, and it is modeled as i.i.d. $\mathcal{CN}(0, \sigma_z^2)$. Then it is possible to obtain a soft estimate of the data symbol $\mathbf{x}_k^{DL}(n)$ at $k$-th MS as:

$$\hat{\mathbf{x}}_k^{DL,UC}(n) = \mathbf{L}_k^H \mathbf{y}_k^{UC}(n). \tag{3.14}$$

Given the above equations, and assuming the use of Gaussian distributed long codewords, it is now possible to write down the achievable rate for the $k$-th user in the UC approach as follows:

$$\mathcal{R}_k = \mathrm{B} \log_2 \mid \mathbf{I} + \mathbf{R}_k^{-1} \mathbf{A}_{k,k} \mathbf{A}_{k,k}^H \mid, \tag{3.15}$$

where

$$\mathbf{A}_{k,l} = \sum_{j \in \mathcal{J}(k)} \sqrt{\eta_{j,l}} \mathbf{L}_k^H \mathbf{H}_{k,j}^H \mathbf{Q}_{l,j} \, . \tag{3.16}$$

and

$$\mathbf{R}_k = \sum_{l \neq k} \mathbf{A}_{k,l} \mathbf{A}_{k,l}^H + \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k \, , \tag{3.17}$$

is the covariance matrix of the interference in the signal received at the $k$-th MS. Now, some algebraic manipulations are needed in order to express the achievable rate $\mathcal{R}_k$ in (3.15) in a form that permits solving the optimization problems considered in Sections 3.6 and 3.7. By letting:

$$\mathbf{B}_{k,l,j} = \mathbf{L}_k^H \mathbf{H}_{k,j}^H \mathbf{Q}_{l,j} \Longrightarrow \mathbf{A}_{k,l} = \sum_{j \in \mathcal{J}(k)} \sqrt{\eta_{j,l}} \mathbf{B}_{k,l,j} \tag{3.18}$$

the covariance matrix (3.17) can be rewritten as

$$\begin{aligned} \mathbf{R}_k = & \sum_{l \neq k} \sum_{j \in \mathcal{J}(l)} \sum_{j' \in \mathcal{J}(l)} \sqrt{\eta_{j,l} \eta_{j',l}} \mathbf{B}_{k,l,j} \mathbf{B}_{k,l,j'}^H + \\ & + \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k \, , \end{aligned} \tag{3.19}$$

thus implying that the rate for the $k$-th MS can be expressed as:

$$\mathcal{R}_k = \mathrm{B} \log_2 \left| \mathbf{I} + \mathbf{R}_k^{-1} \sum_{j,j'} \sqrt{\eta_{j,k} \eta_{j',k}} \mathbf{B}_{k,k,j} \mathbf{B}_{k,k,j'}^H \right| \tag{3.20}$$

Finally, exploiting the fact that the determinant of the product of two matrices factorizes into the product of determinants, we have:

$$\begin{aligned} \mathcal{R}_k = & \underbrace{\mathrm{B} \log_2 \left| \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k + \sum_{l}^{K} \sum_{j} \sum_{j'} \sqrt{\eta_{j,l} \eta_{j',l}} \mathbf{B}_{k,l,j} \mathbf{B}_{k,l,j'}^H \right|}_{g_1(\boldsymbol{\eta})} \\ & - \underbrace{\mathrm{B} \log_2 \left| \sigma_z^2 \mathbf{L}_k^H \mathbf{L}_k + \sum_{l \neq k}^{K} \sum_{j} \sum_{j'} \sqrt{\eta_{j,l} \eta_{j',l}} \mathbf{B}_{k,l,j} \mathbf{B}_{k,l,j'}^H \right|}_{g_2(\boldsymbol{\eta})} \, , \end{aligned} \tag{3.21}$$

where the vector $\boldsymbol{\eta}$ is a compact notation to denote the set of all the downlink

transmit powers $\eta_{j,k}$, for all the values of $k$ and $m$. All previous formulas concern the UC approach but they can be extended to the CF approach with ordinary efforts.

## 3.5 Uplink data transmission

The third phase of the communication protocol amounts to uplink data transmission. We denote by $\mathbf{x}_k^{UL}(n)$ the $P$-dimensional data vector to be transmitted by the $k$-th MS in the $n$-th sample time; the corresponding signal received at the $j$-th AP is expressed as:

$$\mathbf{y}_j(n) = \sum_{k=1}^{K} \sqrt{\widetilde{\eta}_k} \mathbf{H}_{k,j} \mathbf{L}_k \mathbf{x}_k^{UL}(n) + \mathbf{w}_j(n) \tag{3.22}$$

where $\widetilde{\eta}_k = \dfrac{P_{t,k}^{UL}}{\mathrm{tr}(\mathbf{L}_k^H \mathbf{L}_k)}$, and $P_{t,k}^{UL}$ is the uplink transmitted power by the $k$-th MS.

**CF approach**

in this case, each AP forms the following statistic, $\forall k$:

$$\tilde{\mathbf{y}}_{m,k}(n) = \mathbf{Q}_{k,j}^H \mathbf{y}_j(n) \ , \tag{3.23}$$

wherein now the ZF beamformer $\mathbf{Q}_{k,j}$ previously defined is used as a post-coder. Then each AP sends to the CPU the vectors $\tilde{\mathbf{y}}_{m,k}(n)$ via the backhaul link, and the CPU forms the following soft estimate of the data vectors transmitted by th $k$-th MS:

$$\hat{\mathbf{x}}_k^{UL}(n) = \sum_{j=1}^{J} \tilde{\mathbf{y}}_{m,k}(n), \quad k = 1, \ldots, K. \tag{3.24}$$

Plugging (3.22) and (3.23) into (3.24) it can be checked that the signal contributions associated to the data-symbol $\mathbf{x}_k^{UL}(n)$ are coherently summed.

**UC approach**

in this case the signal transmitted by the $k$-th MS is decoded only by the APs belonging to the set $\mathcal{J}(k)$. Accordingly, the CPU performs the following soft estimate:

$$\hat{\mathbf{x}}_k^{UL,UC} = \sum_{j \in \mathcal{J}(k)} \tilde{\mathbf{y}}_{j,k}(n), \quad k = 1, \ldots, K, \tag{3.25}$$

Substituting (3.22) and (3.23) into (3.25), we get:

$$\begin{aligned} \hat{\mathbf{x}}_k^{UL,UC} = & \sum_{j \in \mathcal{J}(k)} \mathbf{Q}_{k,j}^H \sqrt{\tilde{\eta}_k} \mathbf{H}_{k,j} \mathbf{L}_k \mathbf{x}_k^{UL}(n) + \\ & + \sum_{l=1, l \neq k}^{K} \sqrt{\tilde{\eta}_l} \sum_{j \in \mathcal{J}(k)} \mathbf{Q}_{k,j}^H \mathbf{H}_{l,j} \mathbf{L}_j \mathbf{x}_j^{UL}(n) + \\ & + \sum_{j \in \mathcal{J}(k)} \mathbf{Q}_{k,j}^H \mathbf{w}_j(n) \ . \end{aligned} \tag{3.26}$$

Given the above equation, the achievable rate for the $k$-th MS can be easily shown to be expressed in the UC case as

$$\widetilde{\mathcal{R}}_k = B \log_2 \left| \mathbf{I} + \widetilde{\eta}_k \mathbf{R}_k^{-1} \left( \sum_{j \in \mathcal{J}(k)} \mathbf{B}_{k,k,j}^H \right) \left( \sum_{j \in \mathcal{J}(k)} \mathbf{B}_{k,k,j} \right) \right| , \tag{3.27}$$

wherein the interference covariance matrix now is written as

$$\begin{aligned} \mathbf{R}_k = & \ \textstyle\sum_{l \neq k} \widetilde{\eta}_l \Big( \sum_{j \in \mathcal{J}(l)} \mathbf{B}_{l,k,j}^H \Big) \Big( \sum_{j \in \mathcal{J}(l)} \mathbf{B}_{l,k,j} \Big) \\ & + \sigma_w^2 \mathrm{card}(\mathcal{J}(k)) \ , \end{aligned} \tag{3.28}$$

with $\mathrm{card}(\cdot)$ denoting cardinality. The above expression can be used also for the CF case by letting $\mathcal{J}(k) = \{1, 2, \ldots, J\}$, for all $k = 1, \ldots, K$.

## 3.6 Global Energy Efficiency Maximization

In this section we address the problem of power control for energy efficiency maximization considering both the downlink and the uplink channel of the described cellular network. The downlink case is treated in the coming Section 3.6.1, while the uplink scenario is dealt with in Section 3.6.2.

### 3.6.1 Downlink power control

Mathematically the problem is formulated as the optimization program:

$$\max_{\boldsymbol{\eta}} \frac{\sum_{k=1}^{K} \mathcal{R}_k(\boldsymbol{\eta})}{\sum_{j=1}^{J} \left[ \sum_{k \in \mathcal{K}(j)} \delta \eta_{j,k} + P_{c,j} \right]} \tag{3.29a}$$

$$\text{s.t.} \sum_{k \in \mathcal{K}(j)} \eta_{j,k} \leq P_{max,j}, \forall j = 1, \ldots, J \tag{3.29b}$$

$$\eta_{j,k} \geq 0, \forall j = 1, \ldots, J, k = 1, \ldots, K \tag{3.29c}$$

where $P_{c,j} > 0$ is the circuit power consumed at AP $m$, $\delta \geq 1$ is the inverse of the transmit amplifier efficiency, and $P_{max,j}$ is the maximum transmit power from AP $m$; $\boldsymbol{\eta}$ is a $KM \times 1$ vector containing all the transmit power of all AP.

Problem (3.29) is challenging due to its fractional objective, which has a non-concave numerator. This prevents the direct use of standard fractional programming methods such as Dinkelbach's algorithm to solve (3.29) with affordable complexity. In addition, another issue is represented by the large number of optimization variables, i.e. $KM$. To counter both issues, we resort to the successive lower-bound maximization method, introduced in [37][7] (it has already been used in other publications, [38] and [39]), and whose details are summarized in the Appendix A. So, in brief, this method tackles (3.29) by alternatively optimizing the transmit powers of each AP, while keeping the transmit powers of the other APs fixed. However, as it will be shown,

[7] In [37] the method is termed successive upper-bound minimization since there the focus is on minimization problems.

even with respect to the transmit powers of a single AP, the global energy efficiency problem remains challenging and for this reason, each subproblem is tackled by means of sequential optimization, as described next. Consider thus problem (3.29) and define the variable blocks $\boldsymbol{\eta}_j = \left\{\eta_{j,k}\right\}_{k=1,\ldots,K}$, for the CF case, and $\boldsymbol{\eta}_j = \left\{\eta_{j,k}\right\}_{k\in\mathcal{K}(j)}$, for the UC case; the vector $\boldsymbol{\eta}_j$ is $K$-dimensional in the CF case, while in the UC approach it will contain $N$ entries, i.e. the powers to be used to communicate with the MSs in the set $\mathcal{K}(j)$. The global energy efficiency maximization problem with respect to the $j$-th variable block refers to the maximization, with respect to $\boldsymbol{\eta}_j$ of the quantity

$$\frac{\sum\limits_{k=1}^{K}\mathcal{R}_k(\boldsymbol{\eta}_j,\boldsymbol{\eta}_{-j})}{\sum\limits_{j=1}^{J}\left[\sum\limits_{k\in\mathcal{K}(j)}\delta\eta_{j,k}+P_{c,j}\right]}, \tag{3.30}$$

with $\boldsymbol{\eta}_{-j}$ representing the set of all the APs' transmit powers except the ones in $\boldsymbol{\eta}_j$. It can be easily seen that the numerator of the global energy efficiency is not concave even with respect to only the variable block $\boldsymbol{\eta}_j$. In particular, (3.30) has a non-concave numerator, since $\mathcal{R}_k(\boldsymbol{\eta})$, which depends on all transmit powers of the base stations, is the difference of two logarithmic functions with the powers $\boldsymbol{\eta}_j$ of AP $m$ appearing in both logarithmic terms, and the difference of two logarithms is in general not concave. This means that (3.29) can not be solved with affordable complexity by either standard convex optimization nor fractional programming tools. Therefore, plain alternating maximization is not suitable, which motivates the use of the successive lower-bound maximization method. To this end, let us observe that each summand at the numerator of the global energy efficiency can be expressed as the difference between the functions $g_1$ and $g_2$ defined in (3.21). At this point we offer the following result.

**Lemma 1.** *Both functions $g_1$ and $g_2$ in* (3.21) *are concave in $\boldsymbol{\eta}_j$*

*Proof.* As a first step, let us show that the function $f : (x,y) \in \{\mathbb{R}_0^+ \times \mathbb{R}_0^+\} \rightarrow \sqrt{xy} \in \mathbb{R}_0^+$ is jointly concave in $(x,y)$. To this end, the Hessian of $f$ can be

written as

$$\mathcal{H} = \frac{1}{4}\begin{pmatrix} -x^{-3/2}y^{1/2} & x^{-1/2}y^{-1/2} \\ x^{-1/2}y^{-1/2} & -x^{1/2}y^{-3/2} \end{pmatrix}, \tag{3.31}$$

which can be seen to be negative-semidefinite, since its determinant is zero and the elements on the diagonal are non-positive.

Next, we observe that both $g_1$ and $g_2$ in (3.21) are obtained by composing the function $f$ with the $\log_2|(\cdot)|$ function of positive-definite argument. Then, the result follows recalling that the function $\log_2|(\cdot)|$ is matrix-concave and matrix-increasing over the set of positive-semidefinite matrices [40, Section 3]. □

Based on Lemma 1, we argue that (3.21) is the difference between two concave function. Thus, recalling that any concave function is upper-bounded by its first-order Taylor expansion around any given point $\boldsymbol{\eta}_{j,0}$, a concave lower-bound of $\mathcal{R}_k$ can be obtained as:

$$\begin{aligned} \mathcal{R}_k &= g_1(\boldsymbol{\eta}_j) - g_2(\boldsymbol{\eta}_j) \\ &\geq g_1(\boldsymbol{\eta}_j) - g_2(\boldsymbol{\eta}_{j,0}) - \nabla_{\boldsymbol{\eta}_j}^T g_2 \mid_{\boldsymbol{\eta}_{j,0}} (\boldsymbol{\eta} - \boldsymbol{\eta}_{j,0}) \\ &= \overline{\mathcal{R}}_k(\boldsymbol{\eta} - \boldsymbol{\eta}_{j,0}). \end{aligned} \tag{3.32}$$

Then, a suitable approximate problem for the implementation of the sequential optimization method is

$$\max_{\boldsymbol{\eta}_j} \frac{\sum_{k=1}^{K} \overline{\mathcal{R}}_k(\boldsymbol{\eta}_j, \boldsymbol{\eta}_{j,0}, \boldsymbol{\eta}_{-j})}{\delta \sum_{j=1}^{J} \sum_{k\in\mathcal{K}(j)} \eta_{j,k} + P_c} \tag{3.33a}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}(j)} \eta_{j,k} \leq P_{max,j} \tag{3.33b}$$

$$\eta_{j,k} \geq 0,\ k = 1, \ldots, K \tag{3.33c}$$

**Algorithm 2** Sequential algorithm for GEE maximization

set i=0;
choose any feasible $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_J$;
**repeat**
  **for** $j = 1, \ldots, J$ **do**
    **for** $k = 1, \ldots, K$ **do**
      choose any feasible $\boldsymbol{\eta}_{j,0}$;
      Solve (3.33) by Dinkelbach's algorithm and call $\boldsymbol{\eta}_j^*$ the solution;
      update $\boldsymbol{\eta}_{j,0}$: $\boldsymbol{\eta}_{j,0} = \boldsymbol{\eta}_j^*$
    **end for**
  **end for**
**until** convergence

It can be seen that Problem (3.33) fulfills the three properties **P1**, **P2** and **P3** of the sequential method detailed in the Appendix A[8]. Moreover, the numerator of the objective is now concave, which enables the use of fractional programming tools to globally solve (3.33), such as the popular Dinkelbach's algorithm [31]. The overall power control procedure is formulated as in Algorithm 2 and, based on the properties of the successive lower-bound maximization method reviewed in the Appendix A, we can state the following result.

**Proposition 1.** *Algorithm 2 monotonically improves the global energy efficiency value after each iteration and converges to a first-order optimal point of the original global energy efficiency maximization problem in* (3.29).

### An alternative definition for the GEE on the downlink

The definition of the global energy efficiency reported in (3.29) considers a circuit power consumption that does not depend on the transmit power used by each base station. However, a base station that does not transmit will consume a lower circuit power, since it will switch to idle mode. We stress that such a circumstance can arise when maximizing the energy efficiency, since

[8]The first-order Taylor expansion and its derivative coincide with the function and with its derivative when evaluated at the center of the expansion

the marginal increase to the system achievable rate granted by their activation is overweighted by the increased network power consumption. To account for this circumstance, the terms $P_{c,j}$ can be further detailed to depend on the actual used transmit power, namely defining for all $j = 1, \ldots, J$,

$$P_{c,j} = \widetilde{P}_{c,j}(\mathbb{1}[P_T(j) > 0] + 0.5(1 - \mathbb{1}[P_T(j) > 0])) \tag{3.34}$$

where $P_T(j) = \sum_k \eta_{j,k}$ is the power radiated by the $j$-th AP and $\mathbb{1}[P_T(j) > 0]$ is the indicator function of the set $[P_T(j) > 0]$, being 1 when $P_T(j) > 0$ and 0 otherwise. According to (3.34), we assume that the circuit power consumption, equal to $\widetilde{P}_{c,j}$ when the AP is active, is halved when it does not radiate any power.

While this more sophisticated power consumption model accounts for the lower circuit power consumption of idle access points, it leads to a non-differentiable global energy efficiency function, due to the presence of the indicator function I in (3.34). In this case Algorithm 2 is still guaranteed to monotonically increase the global energy efficiency value after each iteration, but no first-order optimality can be guaranteed upon convergence, owing to the non-differentiability of (3.34). Nevertheless, we should remark that it is also possible to approximate the indicator function in (3.34) by a smooth function, such as a sigmoid, thus recovering the first-order optimality property of the algorithm.

### 3.6.2 Uplink power control

With regard to the uplink, the problem of global energy efficiency maximization is formulated as the optimization program:

$$\max_{\boldsymbol{\eta}} \frac{\sum_{k=1}^{K} \widetilde{\mathcal{R}}_k(\widetilde{\eta}_1, \ldots, \widetilde{\eta}_K)}{\left[\sum_{k=1}^{K} \delta \widetilde{\eta}_k + \widetilde{P}_{c,k}\right]} \tag{3.35a}$$

$$\text{s.t.} \quad \widetilde{\eta}_k \text{tr}(\mathbf{L}_k^H \mathbf{L}_k) \leq P_{T,max}, \forall k = 1, \ldots, K \tag{3.35b}$$

$$\widetilde{\eta}_k \geq 0, \forall k = 1, \ldots, K\,, \tag{3.35c}$$

where $P_{T,max}$ denotes the maximum transmit power for the MSs[9]. It is easy to realize that the optimization problem (3.35) has the same structure as (3.29), thus implying that it can be solved by using the same procedure that has been developed for the downlink. It is also worth noting that while in the downlink the number of variables to be optimized is $JK$ in the CF case and $JN$ in the UC case, for the uplink only $K$ transmit powers are to be optimized, thus implying that the power optimization is much less computationally intensive on the uplink than on the downlink.

## 3.7 Sum-rate maximization

While the main focus of this work is on the maximization of the global energy efficiency function, it should be emphasized that the same approach can be adopted also for sum-rate maximization, which can be tackled as a special case of global energy efficiency maximization. Nonetheless, as detailed in the following, there are instances, when perfect channel state information (CSI) is available, in which the sum-rate maximization algorithm reduces to a standard convex optimization problem. Specifically, the global energy efficiency function reduces to the sum-rate function upon plugging $\delta = 0$ and $P_c = 1$ in (3.29a). Thus, in principle, sum-rate maximization can be performed by

[9]For the sake of simplicity this power is assumed to be the same for all the MSs; this assumption however can be easily relaxed.

applying the simplified version of Algorithm 2 wherein standard convex optimization routines can be used to solve Problem (3.33) when $\mu = 0$ and $P_c = 1$ in each iteration[10].

Nevertheless, sum-rate maximization has one peculiarity that should be highlighted with respect to the global energy efficiency maximization scenario. With reference to the case in which perfect CSI is available, it is possible to consider the ZF precoder

$$\mathbf{R}_k = \sigma_z \mathbf{L}_k^H \mathbf{L}_k = \frac{N_{MS}}{P} \sigma_z \mathbf{I} \,, \tag{3.36}$$

thus removing multi-user interference. As a result, the sum-rate function simplifies to:

$$\begin{aligned} \sum_{k=1}^{K} \mathcal{R}_k = \sum_{k=1}^{K} \mathrm{B} \log_2 \Big| \mathbf{I} + \left( \frac{N_{MS}}{P} \sigma_z^2 \mathbf{I} \right)^{-1} \\ \sum_{j,j' \in \mathcal{J}(k)} \sqrt{\eta_{j,k} \eta_{j',k}} \mathbf{C}_{k,j,j'} \Big| \end{aligned} \tag{3.37}$$

where $\mathbf{C}_{k,j,j'}$ is:

$$\mathbf{C}_{k,j,j'} = \mathbf{L}_k^H \mathbf{H}_{k,j}^H \mathbf{Q}_{k,j} \mathbf{Q}_{k,j'}^H \mathbf{H}_{k,j'} \mathbf{L}_k \,, \tag{3.38}$$

which is a concave function by virtue of Lemma 1. In this case, since the sum-rate is already concave and can be globally maximized with polynomial complexity by using standard convex programming theory.

[10]Note that when $\mu = 0$ and $P_c = 1$, (3.33) is no-longer a fractional program.

## 3.8 Performance analysis

### 3.8.1 Simulation setup

We start by defining the used channel model for generating the matrices $\mathbf{H}_{k,j}$. According to the widely used clustered channel model for mmWave frequencies (see [34] and references therein), $\mathbf{H}_{k,j}$ can be expressed as

$$\mathbf{H}_{k,j} = \gamma \sum_{i=1}^{N_{cl}} \sum_{l=1}^{N_{ray}} \alpha_{i,l} \sqrt{L(r_{i,l})} \mathbf{a}_{AP}(\theta_{i,l,k,j}^{AP}) \mathbf{a}_{MS}^{H}(\theta_{i,l,k,j}^{MS}) + \mathbf{H}_{LOS}, \tag{3.39}$$

where $N_{cl}$ is the number of clusters, $N_{ray}$ is the number of rays that we consider for each cluster, $\gamma$ is a normalization factor defined as $\sqrt{\frac{N_{AP}N_{MS}}{N_{cl}N_{ray}}}$, $\mathbf{H}_{LOS}$ is the line-of-sight (LOS) component, $\alpha_{i,l}$ is the complex path gain distributed as $\mathcal{CN}(0,1)$, so that its amplitude is Rayleigh-distributed, $L(r_{i,l})$ is the attenuation related to the path $(i,l)$, $\mathbf{a}_{AP}$ and $\mathbf{a}_{MS}$ are the array responses at the $j$-th AP and at the $k$-th MS, respectively, and they depend on the angles of arrival and departure, $\theta_{i,l,k,j}^{AP}$ and $\theta_{i,l,k,j}^{MS}$, relative to the $(i,l)$-th path of the channel between the $k$-th MS and the $j$-th AP. The path-loss is defined as [41]

$$L(r) = -20\log_{10}\left(\frac{4\pi}{\lambda}\right) - 10n\log_{10}(r) - X_{\sigma}, \tag{3.40}$$

wherein $r$ is the distance between the transmitter and the receiver, $n$ is the path loss exponent, $X_{\sigma}$ is the shadow fading term in logarithmic units with zero mean and $\sigma^2$-variance and $f_0$ is a fixed frequency (see also table 3.1). The $\mathbf{H}_{LOS}$ in (3.39) is written as[11]

$$\mathbf{H}_{LOS} = I(d)\sqrt{N_{AP}N_{MS}}e^{j\eta}\sqrt{L(d)}\mathbf{a}_{AP}(\theta_{LOS}^{AP})\mathbf{a}_{MS}^{H}(\theta_{LOS}^{MS}). \tag{3.41}$$

In the above equation, $\eta \sim \mathcal{U}(0, 2\pi)$, $I(d)$ is a 0-1 random variate indicating if a LOS link exists between the transmitter and the receiver, and $d$ is the

[11] For the ease of notation we omit the subscript $k, m$.

TABLE 3.1: Parameters for Pathloss model

| Scenario | Model Parameters |
|---|---|
| UMi Street Canyon LOS | n=1.98, $\sigma$=3.1 dB |
| UMi Street Canyon NLOS | n=3.19, $\sigma$=8.2 dB |
| UMi Open Square LOS | n=2.89, $\sigma$=7.1 dB |
| UMi Open Square NLOS | n=1.73, $\sigma$=3.02 dB |

transmitter-receiver distance, measured in meters. Denoting by $p$ the probability that $I_{LOS}(d) = 1$, we have, for the UMi (Urban Microcellular) scenarios [42]:

$$p = \min\left(\frac{20}{d}, 1\right)(1 - e^{-\frac{d}{39}}) + e^{-\frac{d}{39}}. \tag{3.42}$$

As for the number of scatterers and their positions, it should be said that, while usually for every AP-MS pair, a random and independently generated set of scatterers is considered to contribute to the channel matrix (3.39). Therefore, in this case, in order to model the possible channel correlation when the devices are closely spaced, we consider the same set of scatterers for the generation of all the channels. In particular, we assume that, in the considered area, there is a given number of random clusters, each one contributing with three rays[12]. Given these clusters, in order to generate the generic channel $\mathbf{H}_{k,j}$ between the $k$-th MS and the $j$-th AP, we consider as active only those clusters falling in an ellipse built around the position of the MS and the AP (see Fig. 3.2). In this way, on one hand we exclude far clusters from contributing to the channel, while, on the other hand, we are guaranteed that devices closely located will have correlated channels since they will be affected by similar sets of scatterers (see Fig. 3.2 for a graphical illustration).

In the following simulation we have considered a carrier frequency of $f_0 = 73\,\text{GHz}$, a bandwidth of $B = 200\,\text{MHz}$ and, with regard to the above channel model, we have simulated the UMi Open Square scenario [42] of size $250 \times 250\,\text{sqm}$. In order to generate the correlated channels, a total of 25.000 randomly deployed clusters (corresponding to a cluster density of 0.4 cluster/sqm.)

[12] As specified in Section 3.8, a density of 0.4 cluster/sqm. will be considered.

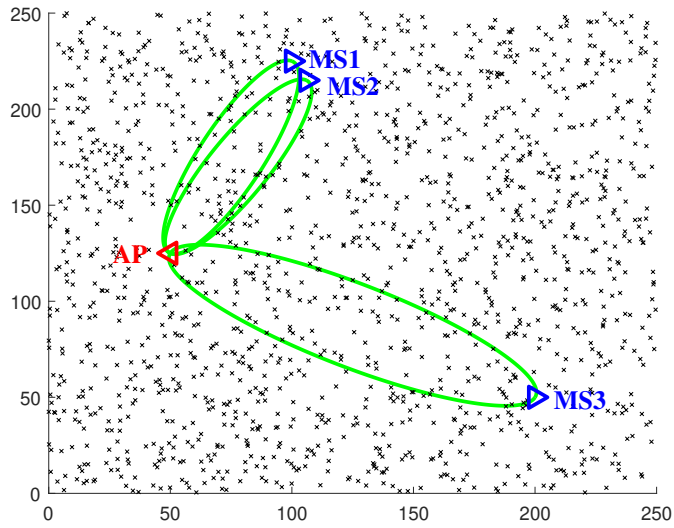



FIGURE 3.2: Detail about the channel generation procedure. For each AP-MS pair, only the scatterers falling into an ellipse built around the locations of the AP and of the MS are considered. This way, we can account for channel correlation when two receivers or two transmitters happen to be closely located. As an example, in the figure, the channels between the AP and MSs 1 and 2 will exhibit some sort of correlation, whereas the channel between the AP and the MS3 is statistically independent from the other channels.

has been generated. The additive white noise at the receiver has a power spectral density of $-174\,\text{dBm/Hz}$ and the receiver noise figure has been set to $F = 6\,\text{dB}$. The simulated system has $M = 80$ APs randomly deployed in the area to cover and equipped with an Uniform Linear Array (ULA) with $N_{AP} = 16$ antennas each. We consider both a scenario with $K = 6$ users, that can be representative of a lightly-loaded network, and a scenario with $K = 16$, that can be representative of a heavily-loaded scenario. We plot results for the UC networking deployment with $N = 1$ and $N = 3$. Each MS is equipped with and ULA with $N_{MS} = 8$ antennas. Although the illustrated algorithms enable each MS to transmit and receive multiple data-streams, a multiplexing order of $P = 1$ is assumed here for the sake of simplicity. The presented results show the global energy efficiency [Mbit/Joule] as defined in (3.29), with the circuit power consumption expressed as in (3.34), and the achievable rate per

user [bit/s], defined as the sum-rate (i.e. the numerator of the global energy efficiency) divided by the number of MSs $K$. For benchmarking purposes, we compare the performance achieved by the proposed power control rules (maximizing the global energy efficiency and the sum-rate) with that achieved by downlink uniform power control, which assumes that $\eta_{j,k} = P_T/K$ in the CF case and that

$$\eta_{j,k} = \begin{cases} \dfrac{P_t}{\mathrm{card}(\mathcal{K}(j))}, & k \in \mathcal{K}(j) \\ 0, & k \notin \mathcal{K}(j) \end{cases}$$

for the UC approach, respectively. The numerical values come from an average over 1000 independent channel scenarios as well as users and access points locations. The convergence criterion for the proposed sequential optimization algorithm is based on the computation of the *relative tolerance*, i.e. the norm of the difference between the current optimized vector and the optimized vector available at the previous iteration, divided by the norm of the current optimized vector. The successive convex approximation routine was directly implemented by the authors with a Matlab script, whereas the routine for maximizing the sum-rate in the case of concave rate was fmincon. All APs have been assumed to have the same maximum feasible transmit power $P_{max}$ on the downlink. The transmit amplifier efficiency of each transmitter has been assumed equal to one, i.e. $\delta = 1$, while the hardware circuit power was modeled according to the model in (3.34) for each AP, with $\widetilde{P}_{c,j} = 1\,\mathrm{W}$, for all $j = 1, \ldots, J$. For the uplink, we instead used $\widetilde{P}_{c,k} = 0.3\,\mathrm{W}$, for all $k = 1, \ldots, K$. In the figures, the dashed line represent the results obtained by means of optimization procedure, instead, the solid line (it has not been reported in the legend in order to avoid a redundancy), with the same color, represent the case of uniform power allocation.

### 3.8.2 Numerical results

Fig. 3.3 compares the global energy efficiency value versus $P_{max}$ achieved by the proposed global energy efficiency-maximizing power control scheme

(labeled as OPT), considering the UC approach in the following scenarios:

- Perfect CSI and FD beamforming.
- Perfect CSI and HY beamforming, with 4 RF chains used in the *BCD-SD* HY beamforming algorithm.
- Imperfect CSI and FD beamforming, with pilot sequences of length $\tau_p = 64$ and uplink transmit power of 100 mW.
- Imperfect CSI and HY beamforming, with pilot sequences of length $\tau_p = 64$ and uplink transmit power of 100 mW, with 4 RF chains used in the *BCD-SD* HY beamforming algorithm.

The plot on the left refers to the UC rule wth $N = 1$, while on the right we have the case $N = 3$. Fig. 3.4 considers the same setting as Fig. 3.3 but reports the downlink average rate per user achieved by the proposed rate-maximizing power control scheme (labeled as OPT). Figs. 3.5 and 3.6 report the same results as Figs. 3.3 and 3.4, respectively, with the difference that they refer to the highly-loaded scenario, i.e. with $K = 16$ users. Inspecting the figures, several considerations can be made. First of all, we see that the proposed power optimization method provides better performance than the uniform power allocation scheme. This is not always true for the case with imperfect CSI (see plot on the top in Fig. 3.5), as a consequence of the fact that, in the imperfect CSI case, the optimization step was performed using the estimated channels, while the plotted curve represents the true global energy efficiency, computed using the real channel coefficients. In other words, as far as the imperfect CSI scenarios are concerned, the metric that is optimized is different from the metric that is plotted. Next, as expected, we notice that FD beamforming and the availability of perfect CSI lead to better performance with respect to the practical situation that imperfect CSI and HY beamforming is to be accounted for. For instance, focusing on the plot on the left in Fig 3.4 and considering a maximum transmit power of 0 dBW, the average rate-per-user drops from 2.65 Gbit/s of the ideal case to 1.65 Gbit/s, for the case in which

both HY beamforming and ICSI are taken into account. Further, when comparing the lightly-loaded scenario with the highly-loaded one, we see that the system global energy efficiency does not dramatically change in the two situations, while, conversely, in a heavily-loaded system with HY beamforming and incomplete CSI the average rate-per-user at 0dBW of maximum transmit power is approximately equal to 500 Mbit/s, with about 60% loss with respect to the average rate-per-user attainable in a lightly loaded scenario. A peculiar behavior that is observed in Figs. 3.4 and 3.6 is that in the case of perfect CSI with FD beamforming with $N = 1$, the uniform and optimal power allocation schemes achieve the same performance. This is due to the fact that the rate is an increasing function of the total available power $P_{max}$ and so, when only one user needs to be served, uniform power allocation coincides with the rate-maximizing power allocation strategy. Instead, this behavior is not observed in Figs. 3.3 and 3.5, where a visible gap is present between the performance with uniform power allocation and with the optimized power allocation also when $N = 1$. This is expected because, unlike the rate, the global energy efficiency is not monotonically increasing with $P_{max}$.

Next, Fig. 3.7 shows the CDFs for the rate-per-user in case of FD and HY beamforming, respectively, considering $P_{max} = 0$ dBW and the lightly loaded scenario. Again the two situations $N = 1$ and $N = 3$ are examined. The curves confirm the findings previously commented. Additionally, it can be seen that for the practical case of incomplete CSI and HY beamforming the scenario with $N = 1$ provides much better performance than the one with $N = 3$ for the vast majority of the users, i.e. it appears to be convenient to have each AP serve just one user than 3 users, except than when considering the unlucky users that happen to be situated on the left tail of the rate-per-user distribution. This behavior can be explained by noticing that in the considered setting, where $M >> K$, the majority of the MSs has in its neighborhood several APs, and letting these APs dedicate their own resources to only MS results in overall increased performance.

Finally, we turn our attention to the performance of the uplink channel. Fig. 3.8 reports the average uplink rate-per-user versus the uplink transmit

power. No power control is performed here. Again, the scenarios $N = 1$ and $N = 3$ are considered. From the figure, it is seen that also for the uplink the rate is a clearly increasing function of the transmit power only in the ideal case of perfect CSI and fully digital beamforming, while, in the other situations, a saturation effect is observed. Results also show that, when focusing on the practical scheme with incomplete CSI and HY beamforming the case $N = 1$ provides better results than the case $N = 3$. In particular, assuming an uplink transmit power of -10dBW a data-rate of about 1.2 Gbit/s can be achieved for the case $N = 1$, versus a data-rate of about 600 Mbit/s for the case $N = 3$. Again, this behavior can be explained by noticing that in a scenario with a dense AP deployment each MS is sorrounded by several APs, and it is better to let each AP to focus its own resourced on just one MS rather then letting the AP to share its resources among multiple MSs.

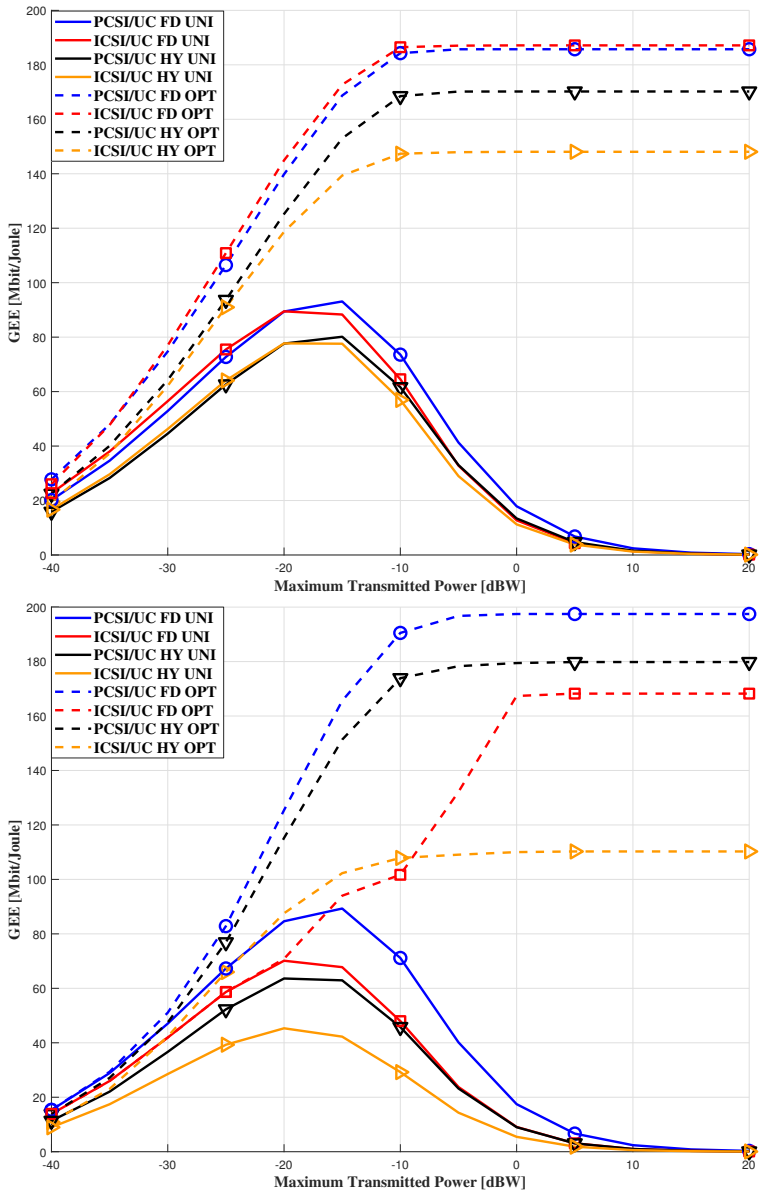



FIGURE 3.3: Global Energy Efficiency with fully-digital (FD) and hybrid beamforming (HY) versus maximum transmit power. On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 6$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$, $\delta = 1$, $P_c = 1$ W.

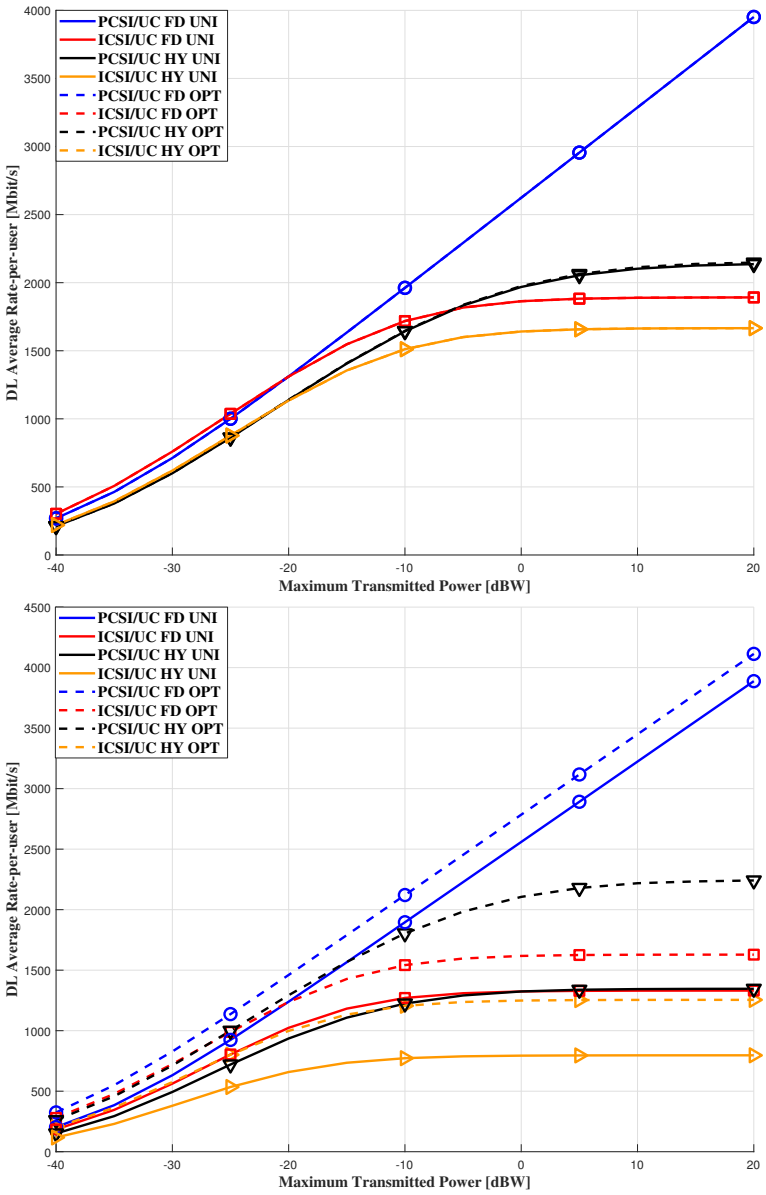



FIGURE 3.4: Average Achievable Rate-per-user with fully-digital (FD) and hybrid beamforming (HY) versus maximum transmit power. On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 6$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$.

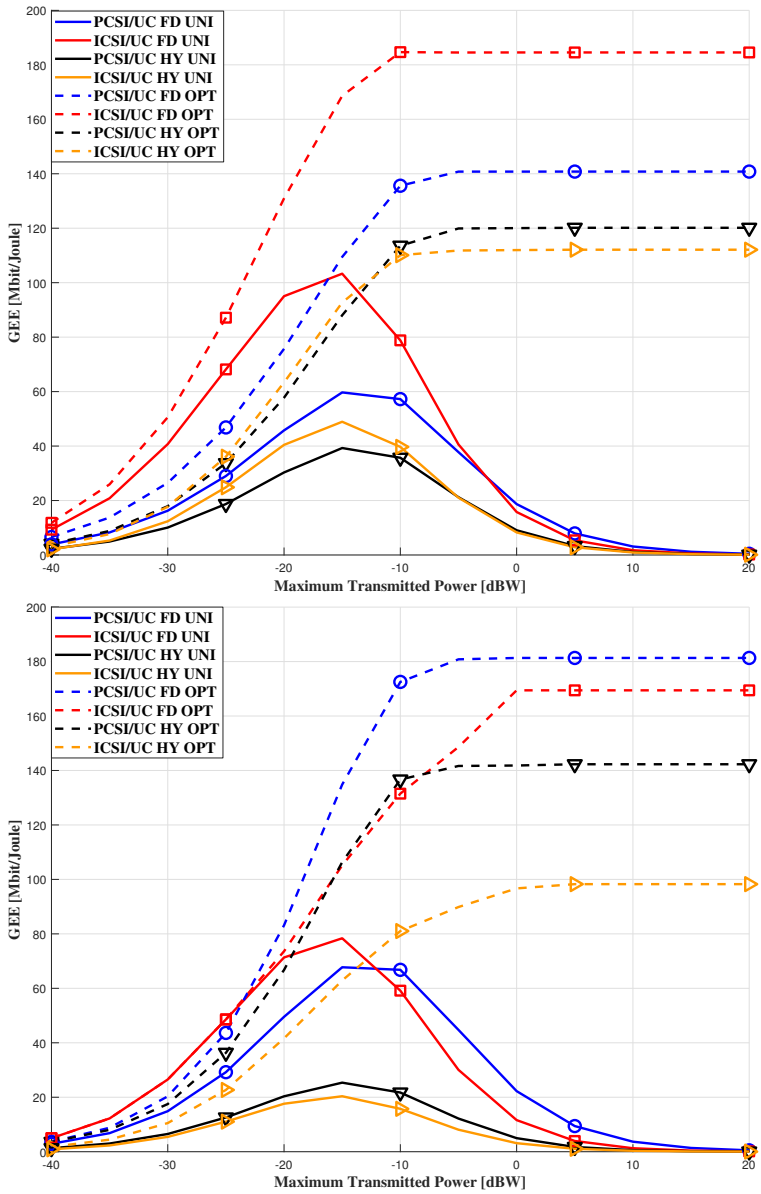



FIGURE 3.5: Global Energy Efficiency with fully-digital (FD) and hybrid beamforming (HY) versus maximum transmit power. On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 16$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$, $\delta = 1$, $P_c = 1$ W.

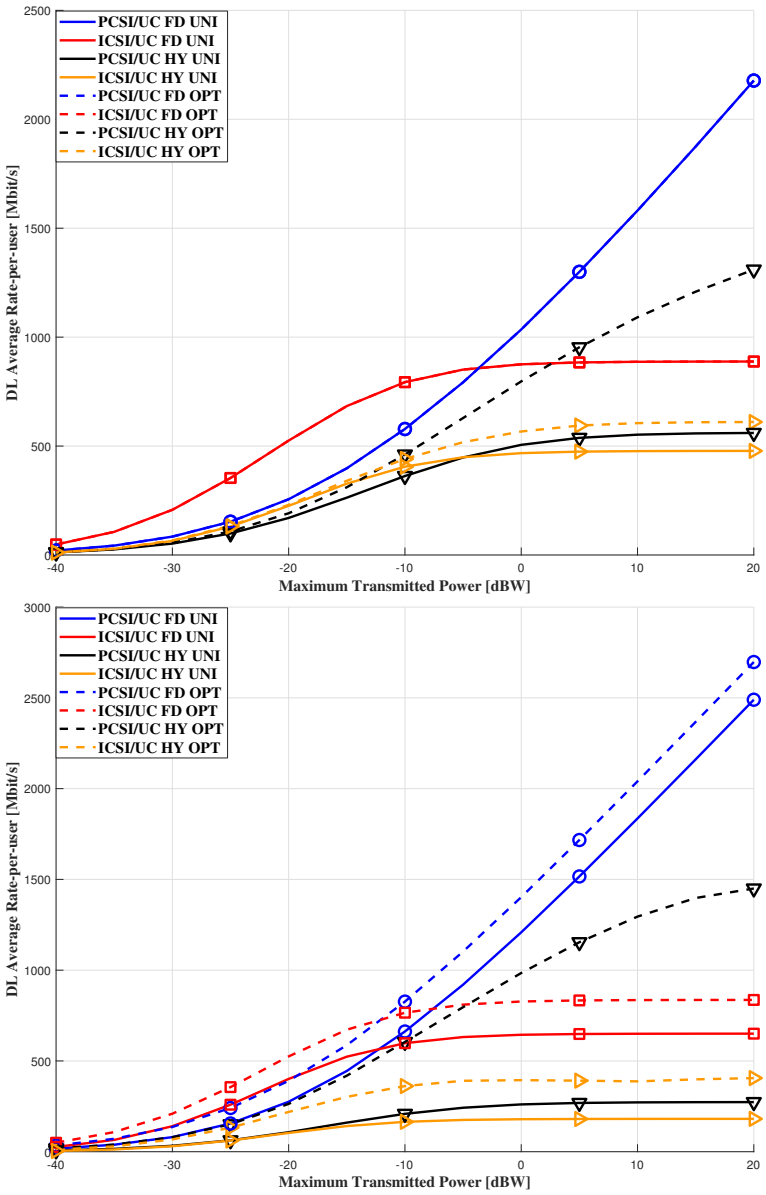



FIGURE 3.6: Average Achievable Rate-per-user with fully-digital (FD) and hybrid beamforming (HY) versus maximum transmit power. On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 16$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$.

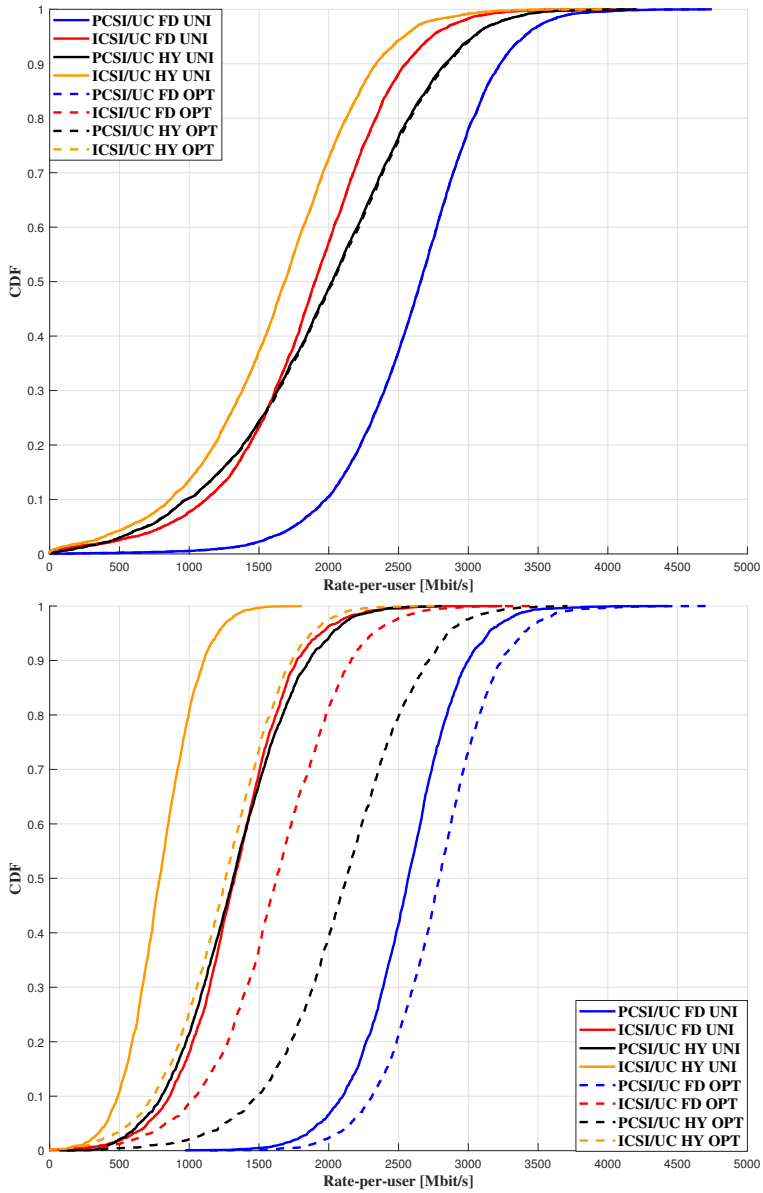



FIGURE 3.7: CDF of rate-per-user with fully-digital (FD) and hybrid beamforming (HY). On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 6$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$.

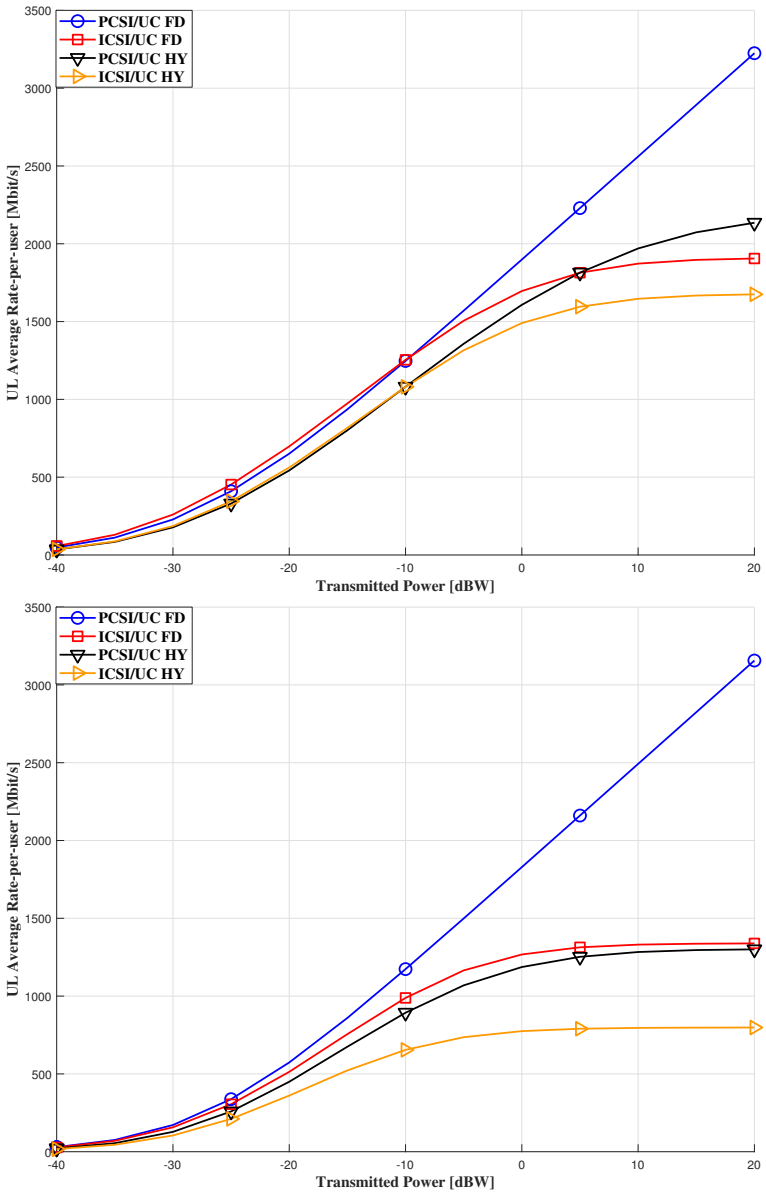



FIGURE 3.8: Uplink Average Achievable Rate-per-user with fully-digital (FD) and hybrid beamforming (HY) versus transmitted power. On the top we have the case N=1 and on the bottom the case N=3. System parameters: $M = 80$, $K = 6$, $N_{AP} \times N_{MS} = 16 \times 8$, $P = 1$.

# Chapter 4

# Cell-Free and Distributed MIMO in Industry 4.0

Referred article "Distributed and Cell-Free Massive MIMO Architectures for URLLC in Indoor Factory Environments" submitted to *IEEE Open Journal of the Communications Society*.

## 4.1 Introduction

In the recent past, the use of wireless communications in factory automation has attracted a lot of interest in the scientific community [43]–[46]. Indeed, using wireless links permits avoiding the installation and maintenance of the network cables, providing increased robusteness and greater flexibility. Wireless factory automation is thus one of the most important use-cases of the fifth-generation (5G) of mobile networks [44]. 5G systems have to satisfy specific requirements such as low latency, high reliability, and availability. This has led to the concept of ultra reliable low latency communications (URLLC) [47], [48], which are instrumental to the development of new applications such as autonomous driving, remote control of drones, and wireless control of actuators (ACs) in factories. To this end, the 3rd Generation Partnership Project (3GPP) has already defined different indoor industrial scenarios, with strict requirements on the latency [49], and several research groups

worldwide are addressing the design of URLLC systems. In [50], the authors addressed the analysis in terms of coverage and capacity for different strategies, wherein a frequency planning is used to enhance the system performance. In [51] the authors show that in industrial scenarios, with focus on URLLC and on traffic analysis, network slicing can be a good strategy to satisfy different requirements given by both low-latency and high data rates. In [52] the authors demonstrate that, in industrial applications, under suitable hypotheses such as the deterministic traffic pattern, it is possible to shorten the size of OFDM preamble to reduce the latency, by adopting techniques for packet detection of OFDM signals. In [53] the authors addressed the problem of resource allocation of short packet transmission for factory scenarios by comparing four different transmission schemes, orthogonal multiple access (OMA), non-orthogonal multiple access (NOMA), relay-assisted transmission and cooperative NOMA (C-NOMA) transmission. Under suitable constraints, the authors described an optimization problem aimed to jointly optimize the blocklength and power allocation to minimize the decoding error probability. In [54], the authors considered a scenario wherein the Access Points (APs) form a distributed multiantenna system, to communicate with the active ACs in the factory. The small scale fading was modeled as the common Rayleigh fading, instead, a path-loss model for industrial scenario was used [55]. They analyzed the performance in term of SINR, for different deployments, different beamforming, in case of equal and optimal power allocation [56]. In [7], the authors carried out the analysis regarding the Distributed MIMO (D-MIMO) systems for industrial scenarios, by exploiting a more realistic 3GPP compliant spatial channel model [57]. The authors taken into account the small scale fading model, in conjunction with a path-loss model defined in [55]. It is shown in the paper that by using a distributed multiantenna architecture it is possible to improve the network reliability. In particular, the authors addressed the problem of Link Adaptation (LA), considering the block error rate (BLER) as KPI and evaluating different transmission modes, beamformers, with varying number of APs and ACs. They focused their attention to a target BLER of $10^{-5}$, which is an extremely low value, required in the

URLLC.

In this work, we study transmission modes in a distributed antenna setting for industrial scenarios. We will exploit a recent network architecture, the Cell-Free (CF) [14]. In CF, all the APs serve all the ACs, and all the APs are connected through a backhaul link to a central processing units (CPU). Nevertheless every AP performs locally channel estimation, and the channel estimates are not shared; also the beamformers are locally evaluated; this leads to a limited usage of the backhaul link. During the downlink data transmission, the CPU sends to the APs the data symbols to be sent to all ACs. In addition to CF, it is possible to derive the User-Centric (UC), similarly to [9], wherein each AC, based on an association rule, has to choose the APs to connect to; this approach permits to reduce the amount of data to be sent through the backhaul link. Then, in case of conjugate beamforming, a lower bound for the Signal-to-Interference-Plus-Noise Ratio (SINR) can be evaluated in a closed form, as in [14]. Moreover, a rate analysis will be addressed, by using the recent Finite Block Length Capacity formula [6]. This formula takes into account the length of the codewords and a target Bit Error Rate (BER). The problem of power control will be also addressed, in order to maximize the minimum SINR across users; this is extremely important when reliability is to be privileged to peak data rates. The Time Division Duplex (TDD) protocol will be considered, in order to exploit the reciprocity of the channel and no channel estimation is required at the ACs.

## 4.2 System model

An industrial scenario represented by an indoor factory hall is considered. We denote by $K$ the total number of ACs equipped with a single antenna element, by $J$ the number of APs distributed in the factory hall, and by $M_{TOT}$ the total number of antenna elements. This implies that $M = M_{TOT}/J$ is the number of antennas for each AP.

We will consider the cases of

a) *centralized deployment*[13], corresponding to the case in which a single AP equipped with $M_{TOT}$ antennas is placed at the center of the factory hall, so that the communication infrastructure resembles a single-cell network;

b) *distributed deployment*, corresponding to the case in which $J$ APs equipped with $M$ antennas each are placed in the factory hall; in this situation the communication infrastructure resembles a multi-cell network. The APs are assumed to be connected to a central processing unit (CPU) through reliable links.

We denote by $\mathbf{h}_{k,j}$ the $(M \times 1)$-dimensional uplink channel vector between the $k$-th AC and the $j$-th AP[14], defined as:

$$\mathbf{h}_{k,j} = \sqrt{\beta_{k,j}}\mathbf{f}_{k,j} \,, \tag{4.1}$$

where $\beta_{k,j}$ represents the large scale fading, i.e. path-loss plus shadowing, while $\mathbf{f}_{k,j}$ is the small scale fading vector, whose entries are a sequence of i.i.d. $\mathcal{CN}(0,1)$ random variates. Since the time division duplex protocol is used, we have channel reciprocity and the downlink channel between the $j$-th AP and the $k$-th AC is expressed as $\mathbf{h}_{k,j}^H$.

### 4.2.1 Transmission modes

Four different transmission modes to serve the ACs are considered in this chapter.

1. *Joint transmission (JT)*. All the APs serve all the ACs by using the same time-frequency resource. In this case, the system is equivalent to a network with a single AP having distributed antennas. All the computations related to the computations of the downlink beamformers are executed in the CPU.

[13]This deployment is actually considered as a benchmark to enable comparison with distributed deployments.

[14]Clearly, for the case of centralized deployment there will be just one channel vector, $\mathbf{h}_{k,1}$, for any $k = 1, \ldots, K$.

2. *Single AP transmission (SAT)*. Each AC $k$ is served only by one AP, whose index is denoted by $j_k$. A reasonable choice is to let every AC to be served by the AP with the strongest large scale fading coefficient, i.e.

$$j_k = \underset{j=1,2,\ldots,J}{\arg\max}\, \beta_{k,j}\,, \;\; k = 1, \ldots, K\,. \tag{4.2}$$

   SAT mode resembles a traditional wireless cellular network where each mobile device is linked to one base station at a time.

3. *Cell-Free transmission (CFT)*. Following the transmission scheme developed in [13], [14], in this case all the APs serve all the users like in the JT case, but the channel estimates computed at the APs are not shared through the CPU and the beamformers are computed locally. The CFT is thus a simplified version of the JT since it requires smaller deployment complexity.

4. *User-Centric transmission (UCT)*. Following the scheme proposed in [18], this transmission mode resembles the CFT, with the difference that now every AC is served by a certain subset of APs. Although several AC-AP association rules can be conceived, in this chapter we will restrict our attention to the case in which every AC is served by a certain number, say $N$, with $N \leq J$, of APs. A possible strategy may be to let the $k$-th AC to be served by the $N$ APs that have the strongest large scale fading coefficients. Otherwise stated, letting $j_k(1), j_k(2), \ldots, j_k(J)$ a sequence of indexes such that $\beta_{k,j_k(1)} \geq \beta_{k,j_k(2)}, \geq \ldots \geq \beta_{k,j_k(J)}$, in the UCT mode the $k$-th AC is served by the APs with indexes $j_k(1), j_k(2), \ldots, j_k(N)$. We will denote by $\mathcal{J}(k)$ the set of APs serving the $k$-th AC. Similarly, $\mathcal{K}(j)$ will denote the set of ACs served by the $j$-th AP. Clearly, if $\widetilde{j} \in \mathcal{J}(\widetilde{k})$, then we have $\widetilde{k} \in \mathcal{K}(\widetilde{j})$.

It is worth noting that in a centralized deployment (i.e. $J = 1$) all the above transmission modes end up coincident. Additionally, it is useful to realize that the transmission modes SAT, CFT and UCT can be treated in an unified way, since they can be obtained through a proper definition of the sets $\{\mathcal{J}(k)\}_{k=1}^{K}$;

in particular, we have that SAT corresponds to the choice $\mathcal{J}(k) = \{j_k(1)\}$; CFT corresponds to the choice $\mathcal{J}(k) = \{1, \ldots, J\}$; and, finally, UCT corresponds to the choice $\mathcal{J}(k) = \{j_k(1), j_k(2), \ldots, j_k(N)\}$. Finally, as regards the load on the backhaul link in distributed deployments, it is easy to realize that JT poses the largest load on the backhaul, followed in decreasing order by CFT, UCT, and SAT.

## 4.3 Transceiver processing

We now describe the needed transceiver processing for the above mentioned transmission modes and deployment strategies.

### 4.3.1 Uplink training

Regardless of the transmission mode and of the deployment strategy, during the uplink training the ACs transmit pilot sequences in order to enable channel estimation at the APs. Let $T$ be the length (in discrete-time samples) of the training sequences. We define by $\mathbf{\Phi} \in \mathcal{C}^{T \times K}$ the matrix containing on its $k$-th column the pilot sequence sent by the $k$-th AC. We assume that this matrix has orthogonal columns, with unitary norm. The received signal, $\mathbf{Y}_j$, at the $j$-th AP is an $M \times T$-dimensional matrix, written as:

$$\mathbf{Y}_j = \sum_{k=1}^{K} \sqrt{\eta_k^{UL-T}} \mathbf{h}_{k,j} \mathbf{\Phi}_{:,k}^H + \mathbf{W}_j \, . \tag{4.3}$$

In the above equation, $\eta_k^{UL-T} = \tilde{\eta}_k^{UL-T} \; T$ is the transmitted power by the $k$-th AC, $\mathbf{\Phi}_{:,k}$ is the $k$-th column of $\mathbf{\Phi}$, $\mathbf{W}_j$ is the $M \times T$-dimensional matrix of thermal noise samples, whose entries are assumed to be i.i.d. $\mathcal{CN}(0, \sigma_w^2)$ random variates. Based on the observable (4.3) at the $j$-th AP, in order to estimate the channel from the $k$-th AC $\mathbf{h}_{k,j}$, a projection onto the $k$-th pilot

sequence is done:

$$\check{\mathbf{y}}_{k,j} = \mathbf{Y}_j \mathbf{\Phi}_{:,k} = \sqrt{\eta_k^{UL-T}} \mathbf{h}_{k,j} + \sum_{l=1, l\neq k}^{K} \sqrt{\eta_l^{UL-T}} \mathbf{h}_{l,j} \mathbf{\Phi}_{:,l}^H \mathbf{\Phi}_{:,k} + \mathbf{W}_j \mathbf{\Phi}_{:,k} \,. \tag{4.4}$$

Given the $M$-dimensional vector $\check{\mathbf{y}}_{k,j}$, and, assuming that the large scale fading coefficients are known at the APs, a linear minimum mean square error (LMMSE) estimate for the channel vector $\mathbf{h}_{k,j}$ is obtained as follows:

$$\hat{\mathbf{h}}_{k,j} = \underbrace{\frac{\sqrt{\eta_k^{UL-T}} \beta_{k,j}}{\sum_{l=1}^{K} \eta_l^{UL-T} \beta_{l,j} \mid \mathbf{\Phi}_{:,l}^H \mathbf{\Phi}_{:,k} \mid^2 + \sigma_w^2}}_{\alpha_{k,j}} \check{\mathbf{y}}_{k,j}. \tag{4.5}$$

Some remarks are now in order. First of all, other channel estimation strategies can be pursued. In this chapter, we will consider LMMSE estimation only due to its simplicity and generally good performance. Next, notice that channel estimation is performed locally at each AP. For CFT and UCT and SAT, this information is exploited locally for computing the beamformer; for the case of JT, instead, these estimates are to be sent to the CPU.

### 4.3.2 Downlink transmission and beamformer choice

Let us detail now the transmitted signal model. Denoting by $x_k(n)$ the unit-energy information symbol to be sent to the $k$-th AC in the $n$-th symbol interval, and by $\eta_{k,j}$ and $\mathbf{q}_{k,j}$ the transmit power and the $M$-dimensional transmit beamformer used at the $j$-th AP to transmit to the $k$-th AC, respectively. Then the signal transmitted by the $j$-th AP for SAT, CFT and UCT can be generally expressed as

$$\mathbf{s}_j(n) = \sum_{k \in \mathcal{K}(j)} \sqrt{\eta_{k,j}} \mathbf{q}_{k,j} x_k(n) \,. \tag{4.6}$$

The signal received by the $k$-th AC can be easily shown to be expressed as

$$\begin{aligned}\hat{x}_k(n) &= \sum_{j=1}^{J} \mathbf{h}_{k,j}^H \mathbf{s}_j(n) + z_k(n) \\ &= \sum_{j\in\mathcal{J}(k)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{k,j} x_k(n) + \sum_{l=1,l\neq k}^{K} \sum_{j\in\mathcal{J}(l)} \sqrt{\eta_{l,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{l,j} x_l(n) + z_k(n),\end{aligned} \tag{4.7}$$

where $z_k(n)$ is the additive thermal noise distributed as $\mathcal{CN}(0,\sigma^2)$.

For the case of JT, instead, denoting by $\eta_k$ the overall power used to transmit to AC $k$, the $M$-dimensional vector transmitted by the $j$-th AP in the $n$-th interval is written as

$$\mathbf{s}_j(n) = \sum_{k=1}^{K} \sqrt{\eta_k}\mathbf{q}_{k,j} x_k(n)\ . \tag{4.8}$$

Upon letting $\mathbf{q}_k = [\mathbf{q}_{k,1}^T, \mathbf{q}_{k,2}^T, \ldots, \mathbf{q}_{k,J}^T]^T$ and $\mathbf{h}_k = [\mathbf{h}_{k,1}^T, \mathbf{h}_{k,2}^T, \ldots, \mathbf{h}_{k,J}^T]^T$ denote the $M_{TOT}$-dimensional overall beamformer and channel vector for AC $k$, the signal received at the $k$-th AC can be shown to be written as

$$\hat{x}_k(n) = \eta_k \mathbf{h}_k^H \mathbf{q}_k x_k(n) + \sum_{l=1,l\neq k}^{K} \eta_l \mathbf{h}_k^H \mathbf{q}_l x_l(n) + z_k(n)\ . \tag{4.9}$$

Deferring to the next section the discussion about the used power control rules, we focus now on the problem of beamforming design. Several beamformers will be considered here.

**Maximum ratio transmission (MRT)**

The MRT beamformer (aka channel matched beamformer) is simply defined as

$$\mathbf{q}_k^{(MRT)} = \frac{\hat{\mathbf{h}}_k}{\|\hat{\mathbf{h}}_k\|}\ , \tag{4.10}$$

for JT mode, and as

$$\mathbf{q}_{k,j}^{(MRT)} = \frac{\hat{\mathbf{h}}_{k,j}}{\|\hat{\mathbf{h}}_{k,j}\|}\ . \tag{4.11}$$

for SAT, CFT, and UCT modes.

**Full zero forcing (FZF)**

As it is well-known, zero-forcing strategies tend to nullify the interference contribution through projection along suitable signal subspaces. For the case of JT, interference cancellation can take place only if $K \leq M_{TOT}$. Under this assumption, letting $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \ldots, \hat{\mathbf{h}}_K]^H$ be the $(K \times M_{TOT})$-dimensional whole estimated channel matrix, and defining $\mathbf{Q} = \hat{\mathbf{H}}^H(\hat{\mathbf{H}}\hat{\mathbf{H}}^H)^{-1}$, the FZF beamformer to be used when transmitting to the $k$-th AC is written as:

$$\mathbf{q}_k^{(FZF)} = \frac{\mathbf{Q}_{:,k}}{\|\mathbf{Q}_{:,k}\|} \,. \tag{4.12}$$

The above beamformer is computed at the CPU; then it is split into $J$ parts with $M$ entries each and sent to the APs.

Let us consider now the case of SAT, CFT and UCT modes. In this case, the zero-forcing projection is to be done on $M$-dimensional vectors, so interference cancellation can be obtained only if $K \leq M$. Under this assumption, letting $\hat{\mathbf{H}}_j = [\hat{\mathbf{h}}_{1,j}, \ldots, \hat{\mathbf{h}}_{K,j}]^H$ be the $(K \times M)$-dimensional whole estimated channel matrix at AP $j$, and defining $\mathbf{Q}_j = \hat{\mathbf{H}}_j^H(\hat{\mathbf{H}}_j\hat{\mathbf{H}}_j^H)^{-1}$, the FZF beamformer to be used at the $j$-th AP when transmitting to the $k$-th AC is written as:

$$\mathbf{q}_{k,j}^{(FZF)} = \frac{\mathbf{Q}_{j_{:,k}}}{\|\mathbf{Q}_{j_{:,k}}\|} \,. \tag{4.13}$$

**Partial zero forcing (PZF)**

As discussed above, FZF beamforming is capable of nullifying the interference only if the number of ACs is not larger than the number of antennas at the APs. Moreover, when the number of ACs approaches the number of antennas at the AP, the performance starts degrading due to the well-known noise enhancement effect [58]. In order to avoid this effect, a compromise solution may be to adopt a PZF beamformer, wherein every AP nulls the interference only to a subset of the ACs in the system. A reasonable choice is to nullify the

interference only towards the ACs to which the greatest harm is produced. In particular, with reference to SAT, CFT and UCT modes, assuming that the $k$-th AC is served by the $j$-th AP (i.e., $k \in \mathcal{K}(j)$), let us denote by $N_j < M$, the number of ACs to be protected from the interference generated by the AP $j$ when transmitting to the $k$-th AC. The PZF beamformer for AP $j$ and AC $k$ can be then obtained as follows. Consider the set $\{\beta_{1,j}, \beta_{2,j}, \ldots, \beta_{K,j}\} - \{\beta_{k,j}\}$ containing $K-1$ coefficients and let $k_j(1), k_j(2), \ldots, k_j(K-1)$ be a sequence of indexes such the $K-1$ coefficients of the set are ordered in decreasing order. Build the $((N_j+1) \times M)$-dimensional matrix $\mathbf{H}_{k,j} = [\hat{\mathbf{h}}_{k,j}, \hat{\mathbf{h}}_{k_j(1),j}, \hat{\mathbf{h}}_{k_j(2),j}, \ldots \hat{\mathbf{h}}_{k_j(N_j),j}]^H$. Define the matrix $\mathbf{Q}_{k,j} = \hat{\mathbf{H}}_{k,j}^H (\hat{\mathbf{H}}_{k,j} \hat{\mathbf{H}}_{k,j}^H)^{-1}$; the PZF beamformer is finally obtained as:

$$\mathbf{q}_{k,j} = \frac{[\mathbf{Q}_{k,j}]_{:,1}}{\|[\mathbf{Q}_{k,j}]_{:,1}\|} \,. \tag{4.14}$$

The PZF beamformer for JT mode can be obtained through similar steps. We omit the details for the sake of brevity.

## 4.4 Performance measures

As performance measures we consider the downlink SINR and the achievable rate under the finite blocklength capacity [6].

### 4.4.1 SINR expressions for SAT, CFT and UCT modes

Given expression (4.7), it is easy to realize that a SINR expression for the $k$-th user is:

$$\mathrm{SINR}_k^{UB} = \frac{\left| \sum_{j \in \mathcal{J}(k)} \sqrt{\eta_{k,j}} \mathbf{h}_{k,j}^H \mathbf{q}_{k,j} \right|^2}{\sigma_k^2 + \sum_{l=1, l \neq k} \left| \sum_{j \in \mathcal{J}(l)} \sqrt{\eta_{l,j}} \mathbf{h}_{k,j}^H \mathbf{q}_{l,j} \right|^2} \,. \tag{4.15}$$

The superscript "UB" stands for "upper-bound". Indeed, since Eq. (4.15) depends on the true channel realizations, using it into the Shannon rate formula $B \log(1 + \mathrm{SINR})$, with $B$ the communication bandwidth, leads to a rate that is

not actually achievable, i.e. to an upper bound of the actual achievable rate [59]. This bound becomes tight when the channel state information accuracy improves, and for perfect channel knowledge it provides the exact value of the achievable rate.

In practical scenarios, however, where channel estimation errors occur, using the SINR expression in (4.15) for computing the achievable rate leads to optimistic results not useful for conservative system design. To circumvent this problem, a bounding technique, known as the "use and then forget" bound [59, Chapter 2], can be used. This bound leads to a SINR expression that when plugged in the achievable rate Shannon formula provides a lower bound for the achievable rate. Assuming that each AC has the knowledge of the channel statistics, starting from (4.7), we can write:

$$\begin{aligned}\hat{x}_k = &\underbrace{\mathbb{E}\left[\sum_{j\in J(k)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{k,j}\right]}_{D_k} x_k^{DL} + \\ &+ \underbrace{\left(\sum_{j\in J(k)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{k,j} - \mathbb{E}\left[\sum_{j\in J(k)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{k,j}\right]\right)}_{B_k} x_k^{DL} + \\ &+ \sum_{l\neq k} \underbrace{\sum_{j\in J(l)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \mathbf{q}_{l,j}}_{I_{k,l}} x_l^{DL} + z_k\end{aligned} \tag{4.16}$$

where $D_k$, $B_k$, and $I_{k,j}$ represent the strength of desired signal, the beamforming gain uncertainty, and the interference caused by the transmission to the $l$-th AC, respectively. A lower bound to the SINR for the $k$-th AC can be thus written as:

$$\mathrm{SINR}_k^{LB} = \frac{|D_k|^2}{\mathbb{E}\left[|B_k|^2\right] + \sum\limits_{l=1,l\neq k}^{K} \mathbb{E}\left[|I_{k,l}|^2\right] + \sigma_k^2} \tag{4.17}$$

The lower bound (4.17) holds for any beamformer. A closed form expression can be easily obtained however only for the case of non-normalized MRT beamforming, i.e. with $\mathbf{q}_{k,j}^{(MRT)} = \hat{\mathbf{h}}_{k,j}$.

In this case, letting $\gamma_{k,j} = \sqrt{\eta_k^{UL-T}}\beta_{k,j}\alpha_{k,j}M$, the SINR lower bound can be shown to be expressed as

$$\mathrm{SINR}_k^{LB} = \frac{\left(\sum\limits_{j\in J(k)} \sqrt{\eta_{k,j}}\gamma_{k,j}\right)^2}{\sum\limits_{l=1}^{K}\sum\limits_{j\in J(l)} \eta_{l,j}\beta_{k,j}\gamma_{l,j} + \sum\limits_{l=1,l\neq k}^{K}\left(\sum\limits_{j\in J(l)} \sqrt{\eta_{l,j}}\frac{\sqrt{\eta_k^{UL-T}}}{\sqrt{\eta_l^{UL-T}}}\frac{\beta_{k,j}}{\beta_{l,j}}\gamma_{l,j}\right)^2 \left|\boldsymbol{\Phi}_{:,k}^H\boldsymbol{\Phi}_{:,l}\right|^2 + \sigma_k^2}. \tag{4.18}$$

The proof of expression (4.18) is reported in Appendix B. Given the fact that non-normalized MRT beamformers are used, the total average power transmitted by AP $j$ is now expressed as

$$P_j = \sum_{k\in\mathcal{K}(j)} \eta_{k,j}\gamma_{k,j}. \tag{4.19}$$

### 4.4.2 SINR expressions for JT mode

Given expression (4.9), and following similar steps as in the previous section, an upper bound for the SINR can be written as:

$$\mathrm{SINR}_k^{UB} = \frac{\eta_k \mid \mathbf{h}_k^H\mathbf{q}_k \mid^2}{\sigma_k^2 + \sum\limits_{l=1,l\neq k}^{K} \eta_l \mid \mathbf{h}_k^H\mathbf{q}_l \mid^2}, \tag{4.20}$$

where $\sigma_k^2$ is the thermal noise variance.

Similarly, a lower bound can be also derived through the "use and then forget" bound. Indeed, starting from (4.9), we have:

$$\hat{x}_k^{UC} = \eta_k(\underbrace{\mathbb{E}[\mathbf{h}_k^H\mathbf{q}_k]}_{D_k})x_k^{DL} + \eta_k(\underbrace{\mathbf{h}_k^H\mathbf{q}_k - \mathbb{E}[\mathbf{h}_k^H\mathbf{q}_k]}_{B_k})x_k^{DL} + \sum_{l\neq k}\eta_l\underbrace{\mathbf{h}_k^H\mathbf{q}_l}_{I_{k,l}}x_l^{DL} + z_k\,, \tag{4.21}$$

which leads to the following expression

$$\mathrm{SINR}_k^{LB} = \frac{\eta_k |D_k|^2}{\eta_k \mathbb{E}\left[|B_k|^2\right] + \sum\limits_{l=1, l\neq k}^{K} \eta_l \mathbb{E}\left[|I_{k,l}|^2\right] + \sigma_k^2}. \tag{4.22}$$

Expression (4.22) can be given an explicit expression when non-normalized MRT beamforming is used, i.e. when $\mathbf{q}_{k,j} = \hat{\mathbf{h}}_{k,j}$. As fully detailed in Appendix C, we thus have

$$\mathrm{SINR}_k^{LB} = \frac{\eta_k \gamma_k^2}{\sum\limits_{l=1}^{K} \eta_l \sqrt{\eta_l^{UL-T}} \mathrm{tr}(\mathbf{B}_l \mathbf{D}_l^H \mathbf{B}_k) + \sum\limits_{l=1, l\neq k}^{K} \eta_l \eta_k^{UL-T} tr^2(\mathbf{D}_l \mathbf{B}_k) |\mathbf{\Phi}_{:,k}^H \mathbf{\Phi}_{:,l}|^2 + \sigma_k^2}, \tag{4.23}$$

where $\mathbf{D}_k = \mathrm{blkdiag}\left(\alpha_{k,1}\mathbf{I}_M, \alpha_{k,2}\mathbf{I}_M, \ldots \alpha_{k,J}\mathbf{I}_M\right)$ and $\mathbf{B}_k = \mathrm{blkdiag}\left(\beta_{k,1}\mathbf{I}_M, \beta_{k,2}\mathbf{I}_M, \ldots \beta_{k,J}\mathbf{I}_M\right)$. The total average transmitted power is now expressed as $\sum\limits_{l=1}^{K} \eta_l \gamma_l$.

### 4.4.3 Finite Block Length Capacity

The SINR expression derived above can be used to compute upper and lower bounds to the achievable rate. Traditionally, the rate expression is written as $\mathrm{R}_k = B \log_2(1 + \mathrm{SINR}_k)$ where $B$ is the bandwidth, but this formula holds under the assumption that very long, Gaussian-distributed, codewords are employed. When latency constraints are active, as usually happens in the case of remote machinery control, coding schemes with short codewords are used and the traditional rate formula does not hold any longer. In [6], the finite block-length capacity formula was derived; in particular, letting $m$ be the length in symbols of the used codeword and $\epsilon$ the target packet-error-rate, the $k$-th AC rate in the finite blocklength regime can be approximated as

$$\mathrm{R}_k \approx B\left[\log_2(1 + \mathrm{SINR}_k) - \sqrt{\frac{1}{2m}\left(1 - \frac{1}{(1+\mathrm{SINR}_k)^2}\right)} \frac{\mathrm{Q}^{-1}(\epsilon)}{\ln(2)}\right], \tag{4.24}$$

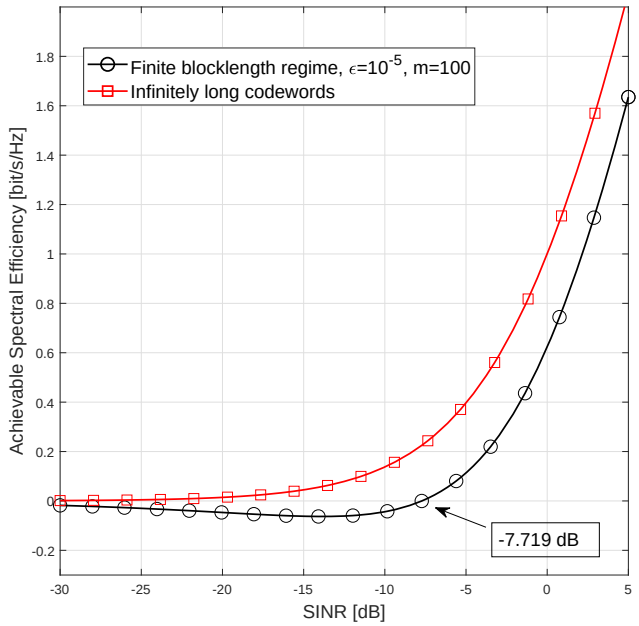



FIGURE 4.1: Achievable spectral efficiency versus SINR for the case of finite blocklength regime and of infinitely long codewords.

with $\mathrm{Q}^{-1}(\cdot)$ denoting the inverse Q-function. Fig. 4.1 reports the achievable spectral efficiency (that is obtained by dividing the rate by the communication bandwidth $B$), for the case of finite blocklength and of infinitely long codewords versus the SINR, assuming $\epsilon = 10^{-5}$ and $m = 100$. As expected, the achievable spectral efficiency for the case of finite blocklength is smaller than that for the case of infinitely long codewords. Moreover, for SINR values smaller than $-7.719$ dB, Eq. (4.24) provides negative values: this means that for such values of SINR it is not possible to communicate reliably given the chosen values of $\epsilon$ and $m$, and, thus, the corresponding rate must be set equal to zero.

## 4.5 Power control rules

### 4.5.1 Uniform power allocation

The simplest form of power control is the uniform power allocation (UPA), wherein each AP uniformly splits its available power budget across the ACs it is connected to. In particular, for the case of SAT, CFT and UCT modes, denoting by $P_{\max,j}$ the maximum power available at AP $j$, the power coefficients $\eta_{k,j}$ are expressed as

$$\eta_{k,j} = \begin{cases} \frac{P_{\max,j}}{|\mathcal{K}(j)|} & k \in \mathcal{K}(j), \\ 0 & k \notin \mathcal{K}(j), \end{cases} \tag{4.25}$$

for the case in which normalized beamformers are used. If, instead, a non-normalized MRT beamformer is used (this is indeed needed in order to be able to obtain the lower bound SINR expression in (4.18)), then we have:

$$\eta_{k,j} = \begin{cases} \frac{P_{\max,j}}{|\mathcal{K}(j)|\gamma_{k,j}} & k \in \mathcal{K}(j), \\ 0 & k \notin \mathcal{K}(j)\ ;. \end{cases} \tag{4.26}$$

For the case of JT, instead, denoting by $P_{\max}$ the maximum available power, and assuming that normalized beamformers are used, we simply have $\eta_k = P_{\max}/K$, $\forall k = 1, 2, \ldots, K$. When a non-normalized MRT beamformer is used, we have instead $\eta_k = \frac{P_{\max}}{K\gamma_k}$.

### 4.5.2 Min-SINR maximum power allocation

In an industrial environment with wireless control of the machinery, it is of the utmost importance to ensure that all the ACs can be properly controlled. Optimizing thus the performance of the ACs with the most unfavorable channels is thus a crucial task. To this end, it may be helpful to allocate transmit power in order to maximize the minimum SINR across the $K$ ACs. Notice that, since

the achievable rate is an increasing function of the SINR[15], this strategy also maximizes the minimum of the ACs' rates.

**SAT, CFT and UCT modes**

For SAT, CFT, and UCT modes, the $k$-th AC SINR expression (4.18) can be written as

$$\mathrm{SINR}_k = \frac{\left( \sum\limits_{j\in J(k)} \sqrt{\eta_{k,j}} a_{k,j} \right)^2}{\sum\limits_{l=1}^{K} \sum\limits_{j\in J(l)} \eta_{l,j} b_{l,j}^{(k)} + \sum\limits_{l=1, l\neq k}^{K} \left( \sum\limits_{j\in J(l)} \sqrt{\eta_{l,j}} c_{l,j}^{(k)} \right)^2 \varphi_{k,l} + \sigma_k^2} . \tag{4.27}$$

with suitable definitions of the coefficients $a_{k,j}$, $b_{l,j}^{(k)}$ and $c_{l,j}^{(k)}$, and $\varphi_{k,l} = \left| \mathbf{\Phi}_{:,k}^H \mathbf{\Phi}_{:,l} \right|^2$
The $k$-th AC SINR expression in (4.15) can be also represented as in (4.27) assuming $c_{l,j}^{(k)} = 0 \ \forall l, j, k$, and $b_{l,j}^{(k)} = 0$ when $l = k$. Let us focus for the moment on the SINR lower bound expression in (4.18).
Letting $\mathcal{P} = \left\{ \eta_{k,j} \ : \ j \in \mathcal{J}(k), \forall k = 1, \ldots, K \right\}$ be the set of unknown powers to be optimized, the following problem can be considered

$$\begin{cases} \max\limits_{\mathcal{P}} \min\limits_{k=1,\ldots,K} \dfrac{\left( \sum\limits_{j\in J(k)} \sqrt{\eta_{k,j}} a_{k,j} \right)^2}{\sum\limits_{l=1}^{K} \sum\limits_{j\in J(l)} \eta_{l,j} b_{l,j}^{(k)} + \sum\limits_{l=1, l\neq k}^{K} \left( \sum\limits_{j\in J(l)} \sqrt{\eta_{l,j}} c_{l,j}^{(k)} \right)^2 \varphi_{k,l} + \sigma_k^2} , \\ \text{subject to:} \begin{cases} \eta_{k,j} \geq 0 \ , \ k = 1, \ldots, K \ \ j = 1, \ldots, J \ , \\ \sum\limits_{k\in\mathcal{K}(j)} \eta_{k,j} \gamma_{k,j} \leq P_{\max,j} \ , \ \ j = 1, \ldots, J \ . \end{cases} \end{cases} \tag{4.28}$$

Introducing the new variables $p_{k,j} \triangleq \sqrt{\eta_{k,j}}$, $\epsilon_{l,k} \triangleq \left( \sum\limits_{j\in J(l)} p_{l,j} c_{l,j}^{(k)} \right)^2$, and $\delta_j^{(k)} \triangleq \sum\limits_{l\in\mathcal{K}(j)} p_{l,j}^2 b_{l,j}^{(k)}$, and letting $\mathcal{Q} = \left\{ \eta_{k,j} \ : \ j \in \mathcal{J}(k), \forall k = 1, \ldots, K \right\} \cup \left\{ \delta_j^{(k)} \ : \ j = 1, \ldots, J \right\} \cup$

[15]This is also true for the rate expression in (4.24), provided that the negative values of the rates in the low SINR region are substituted by zero.

$\{\epsilon_{l,k} \,:\, l,k \in \{1,\dots,K\}^2,\, l \neq k\}$ the new set of variables to be optimized, problem (4.28) can be equivalently re-written as[16]

$$\begin{cases} \max\limits_{\mathcal{Q}} \min\limits_{k=1,\dots,K} \dfrac{\left(\sum\limits_{j\in J(k)} p_{k,j} a_{k,j}\right)^2}{\sum\limits_{j=1}^{J} \delta_j^{(k)} + \sum\limits_{l=1,l\neq k}^{K} \epsilon_{l,k}\varphi_{k,l} + \sigma_k^2}, \\ \text{subject to:} \begin{cases} p_{k,j} \geq 0\,,\ k=1,\dots,K\ \ j=1,\dots,J\,, \\ \sum\limits_{k\in\mathcal{K}(j)} p_{k,j}^2 \gamma_{k,j} \leq P_{\max,j}\,,\ j=1,\dots,J\,, \\ \left(\sum\limits_{j\in J(l)} p_{l,j} c_{l,j}^{(k)}\right)^2 \leq \epsilon_{l,k}\,, \forall l \neq k\,, \\ \sum\limits_{l\in\mathcal{K}(j)} p_{l,j}^2 b_{l,j}^{(k)} \leq \delta_j^{(k)}\,,\ j=1,\dots,J\,. \end{cases} \end{cases} \tag{4.29}$$

Now, it is easy to realize that the objective function in (4.29) is quasi-concave. To see this, notice that for any $t > 0$, the upper level set described by the inequality

$$\frac{\left(\sum\limits_{j\in J(k)} p_{k,j} a_{k,j}\right)^2}{\sum\limits_{j=1}^{J} \delta_j^{(k)} + \sum\limits_{l=1,l\neq k}^{K} \epsilon_{l,k}\varphi_{k,l} + \sigma_k^2} \geq t \tag{4.30}$$

can be rewritten as

$$\frac{1}{\sqrt{t}}\left(\sum_{j\in J(k)} p_{k,j} a_{k,j}\right) \geq \left[\sum_{j=1}^{J} \delta_j^{(k)} + \sum_{l=1,l\neq k}^{K} \epsilon_{l,k}\varphi_{k,l} + \sigma_k^2\right]^2 \tag{4.31}$$

[16] Problem (4.29) is equivalent to (4.28) since at the optimal point the third and fourth constraints in (4.29) are satisfied with equality. The proof of this statement can be obtained by contradiction and is omitted for the sake of brevity.

The right-hand-side of inequality (4.31) can be interpreted as the norm of the vector

$$\mathbf{s}_k \triangleq \left[ \sqrt{\delta_1^{(k)}}. \ldots, \sqrt{\delta_J^{(k)}}, \underbrace{\sqrt{\epsilon_{1,k}\varphi_{k,1}}, \sqrt{\epsilon_{2,k}\varphi_{k,2}}, \ldots, \sqrt{\epsilon_{K,k}\varphi_{k,K}}}_{k\text{-th term not present}}, \sigma_k \right]^T, \quad (4.32)$$

and, thus, the constraint (4.31) is the second-order cone $\|\mathbf{s}_k\| \leq \frac{1}{\sqrt{t}} \left( \sum_{j \in J(k)} p_{k,j} a_{k,j} \right)$. Based on the above arguments, it this follows that the optimal solution to (4.29) can be obtained by solving a sequence of convex feasibility programs, as detailed in Algorithm 3.

The case in which the SINR expression to be considered in the optimization problem is the upper bound (4.15) can be treated similarly, with the only difference that the maximum power constraint is now written as $\sum_{k \in \mathcal{K}(j)} p_{k,j}^2 \leq P_{\max,j}$. We omit the details for the sake of brevity.

### JT mode

For the JT mode, it is seen that both the $k$-th AC SINR expressions (4.20) and (4.23) can be written as

$$\text{SINR}_k = \frac{\eta_k a_k}{\sum_{j=1}^{K} \eta_j b_{k,j} + \sigma_k^2}, \qquad k = 1, \ldots, K \quad (4.34)$$

**Algorithm 3** Bisection algorithm for solving problem (4.28)

1: Choose $t_{\min} = 0$ and $t_{\max}$ as a number certainly larger than the maximun SINR, e.g. $t_{\max} = 10^8$. Choose a tolerance $\epsilon > 0$.
2: **while** $t_{\max} - t_{\min} < \epsilon$ **do**
3: Set $t = \frac{t_{\max}+t_{\min}}{2}$
4: Solve the following convex feasibility program

$$\begin{aligned}
&\|\mathbf{s}_k\| \leq \frac{1}{\sqrt{t}}\left(\sum_{j\in J(k)} p_{k,j}a_{k,j}\right), k = 1,\ldots,K \\
&p_{k,j} \geq 0\,,\ k = 1,\ldots,K\ \ j = 1,\ldots,J\,, \\
&\sum_{k\in\mathcal{K}(j)} p_{k,j}^2\gamma_{k,j} \leq P_{\max,j}\,,\ j = 1,\ldots,J\,, \\
&\left(\sum_{j\in J(l)} p_{l,j}c_{l,j}^{(k)}\right)^2 \leq \epsilon_{l,k}\,, \forall l \neq k\,, \\
&\sum_{l\in\mathcal{K}(j)} p_{l,j}^2 b_{l,j}^{(k)} \leq \delta_j^{(k)}\,,\ j = 1,\ldots,J\,.
\end{aligned} \tag{4.33}$$

5: **if** Problem (4.33) is feasible **then**
6: $t_{\min} = t$
7: **else**
8: $t_{\max} = t$
9: **end if**
10: **end while**

with suitable definitions for the coefficients $\{a_k\}_{k=1}^K$ and $\{b_{k,j}\}_{k,j=1}^K$. Letting $\eta = [\eta_1, \eta_2, \ldots, \eta_K]^T$ be the vector power to be optimized, the following problem is to be considered:

$$\begin{cases}
\max\limits_{\mathbf{J}} \min\limits_{k=1,\ldots,K} \dfrac{\eta_k a_k}{\sum\limits_{j=1}^K \eta_j b_{k,j} + b_{0,k}}\,, \\
\text{subject to:} \begin{cases} \eta_k > 0\,,\ k = 1,\ldots,K\,, \\ \sum\limits_{l=1}^K \eta_l \leq P_{\max}\,. \end{cases}
\end{cases} \tag{4.35}$$

**Algorithm 4** Bisection algorithm for solving problem (4.35)

Choose $t_{\min} = 0$ and $t_{\max}$ as a number certainly larger than the maximun SINR, e.g. $t_{\max} = 10^8$. Choose a tolerance $\epsilon > 0$.
**while** $t_{\max} - t_{\min} < \epsilon$ **do**
  Set $t = \frac{t_{\max}+t_{\min}}{2}$
  Solve the convex feasibility program (4.36)
  **if** Problem (4.36) is feasible **then**
    $t_{\min} = t$
  **else**
    $t_{\max} = t$
  **end if**
**end while**

Problem (4.35) can be equivalently re-written as

$$\begin{cases} \max\limits_{\boldsymbol{\eta},t} t\,, \\ \text{subject to:} \begin{cases} \dfrac{\eta_k a_k}{\sum\limits_{j=1}^{K} \eta_j b_{k,j} + b_{0,k}} \geq t\,,\ k = 1,\ldots,K\,, \\ \eta_k > 0\,,\ k = 1,\ldots,K\,, \\ \sum\limits_{l=1}^{K} \eta_l \leq P_{\max}\,. \end{cases} \end{cases} \tag{4.36}$$

Problem (4.36) can be solved efficiently by a bisection search, in each step solving a sequence of convex feasibility problems [40] as detailed in Algorithm 4. In problems (4.35) and (4.36), the maximum power constraints $\sum_{l=1}^{K} \eta_l \leq P_{\max}$ refers to the case in which normalized beamformers are used; if, instead non-normalized MRT beamformers are employed, then the constraint becomes $\sum_{l=1}^{K} \eta_l \gamma_l \leq P_{\max}$.

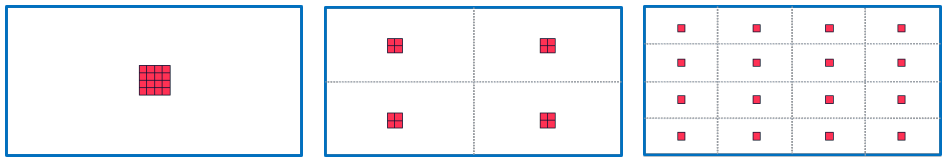
## 4.6 Numerical Results

### 4.6.1 Simulation setup

We assume an industrial scenario with a factory hall of dimensions $100 \times 50 \times 6$ m$^3$; we consider a system bandwidth of 100 MHz at the carrier frequency of 3.7 GHz. We denote by $K = 16$ the total number of ACs equipped with a single antenna element with an height of 1.5 m, and by $M_{TOT}$ the total number of antennas to be subdivided among all the APs. The quantity $J = 4$ denotes the number of APs distributed in the factory hall (see Fig.4.2, wherein different deployments are illustrated) and $M = M_{TOT}/J$ is the number of antennas for each AP. The APs are placed at a height of 6 meters. In the following, numerical results are shown in terms of the CDF of the AC's SINR, and in terms of achievable rate-per-user versus the error probability. The latter curve has been obtained by choosing a value for the error probability $\epsilon$ and by computing the average finite blocklength rate in (4.24) with $m = 100$.

### 4.6.2 Path-loss model

We will focus our attention at the downlink (DL) transmission, and we assume that the total transmit power is 20 dBm. We consider an additive white Gaussian noise (AWGN) process with power spectral density of -174 dBm/Hz and a receiver noise figure (NF) of 7 dB, and a bandwidth of 100 MHz. Regarding the channel model, our starting point are the 3GPP Indoor Office (InO) and Indoor Mixed (InM) models [57]. Furthermore, we modify them considering the proposal in [55], where models for path-loss, shadowing, and line of sight (LOS) probability have been proposed based on extensive measurements done in two different operational factories at 3.5 GHz. Such novel Indoor Industrial (InI) model considers different deployments, distinguishing i) elevated (El) from clutter embedded (Cl) APs and ii) open (Op) from dense (De) production spaces, proposing for each configuration specific values of the aforementioned large scale fading parameters; all the details are reported in [55, Tab. 3].



(A) Centralized (B) Partially distributed (C) Fully distributed

FIGURE 4.2: Deployments

First, we compare in Fig. 4.3 the InI models proposed in [55] against the InO and InM developed in 3GPP for indoor scenarios by showing the cumulative distribution function (CDF) of the SINR achieved by JT with $J = 16$ APs serving $K = 1$ AC with MRT and with perfect channel knowledge. We observe that the achieved SINR is quite large, as just one AC is active (i.e., there is no interference) on the whole factory floor, and a fully distributed deployment is used. However, there is significant difference among the different models, and it is seen that adopting the InI-Cl-De the value of the SINR CDF at $10^{-3}$ is almost 20 dB lower when compared to InO. Based on such result, in the rest of the numerical results we will consider then the InI-Cl-De model as it represents the most challenging scenario and we will provide results only considering imperfect channel knowledge for beamforming design as described in Section 4.3.1 with pilot length $T = 16$.

### 4.6.3 Comparing UP and LB SINR behaviour

Fig. 4.4 shows the gap between the Upper Bound and the Lower Bound, in term of SINR and of average rate per user versus error probability, when the MRT beamformer is used, for the JT, CFT and SAT transmission modes. We see that the gap between the bounds is rather limited, and this is of course a positive result since the bounds can be assumed to be representative of the true SINR values with a limited approximation error. The gap in case of JT is a bit larger than that of the other transmission modes, i.e. about 1 dB, compared to $\sim$0.5-0.7 dB in case of CFT and SAT, in the low SINR region. Based on the evidence that there is a limited gap between the UB and LB, in the sequel we will report the UP curves only for the sake of clarity and plot readability.

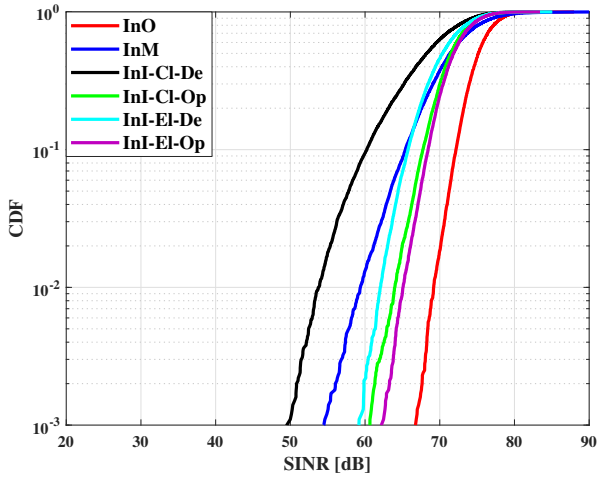



FIGURE 4.3: CDF of the SINR when using different channel models with MRT, JT, $J = 16$, and $K = 1$.

### 4.6.4 Impact of beamformers

Fig. 4.5 presents the achieved performance when MRT and FZF beamformers are employed, for different transmission modes, i.e. JT, CFT, and SAT. Inspecting the curves, an interesting trade-off emerges. Focusing on the plot of the SINR CDF, it is seen that for JT the ZF strategy is uniformly better than the MRT strategy. Conversely, for CFT and SAT the curves corresponding to ZF and MRT cross at some point. So, as far as reliability is concerned, i.e. in the lower-left part of the curve, the MRT achieves better performance than the ZF. This may be explained by noticing that users with weak channels, are mostly endangered by noise rather than by interference, and, so the suffer the noise enhancement effect that is produced by the Zf processing. Results also show that in the lower-left part of the curve CFT provides better performance than SAT, while the opposite happens in the upper-right part of the CDF. If we look at the average rates, instead (i.e., the subplot on the right), the ZF has always better performance than the MRT beamformer, and SAT outperforms CFT. Fig. 4.6 provides a comparison between the ZF beamformer and the PZF (evaluated with $N_j = 7$ and $N_j = 8$), for CFT and SAT. Again, the CDF curves cross at some point. In the lower-left part of the plot, PZF with

$N_j = 7$ exhibits the best performance. In terms of mean rates, CFT with PZF and $N_j = 8$ achieves the best performance, with more than 20% gains over the ZF beamformer.

### 4.6.5 Advantages of UC processing

In Fig. 4.7 the UC transmission mode is contrasted with CFT. Results clearly show that UC transmission outperforms, both in the lower-left part of the plot reporting the CDF of the SINR, and in terms of mean rates, the other transmission modes. Indeed, UC retains the advantages of distributed antenna transmissions, but, at the same time, avoinds the inefficient situation where APs have to waste their energy to transmit to far users with weak channel coefficients. Comparing the results in Fig. 4.7 with those in Fig. 4.5 it appears that JT outperforms the UC transmission mode, but this was to be expected since JT performs a highly complex joint processing

### 4.6.6 Impact of power control

In Fig. 4.8 we finally assess the effect of the power control rule, assuming ZF beamforming, and comparing SAT, JT, CFT and UC transmissionj mode with $N = 3$. Results show that the power control rule is highly effective in the lower-left part of the CDF of the SINR curve. Again, results confirm that the UCT mode is outperformed by the JT only, while outperforming the SAT and CFT modes.

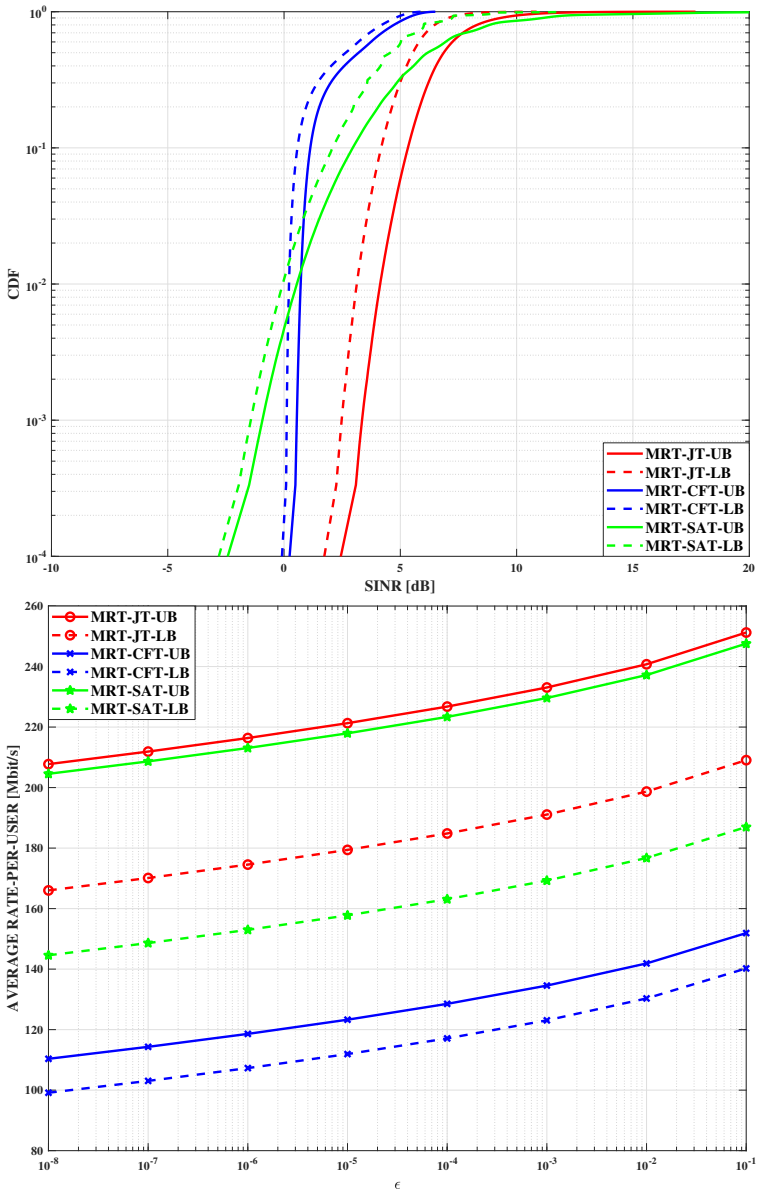



FIGURE 4.4: On the top we have the SINR CDF and on the bottom the average rate-per-user with MRT for different transmission modes, JT, CFT, and SAT; comparison between Upper Bound and Lower Bound.

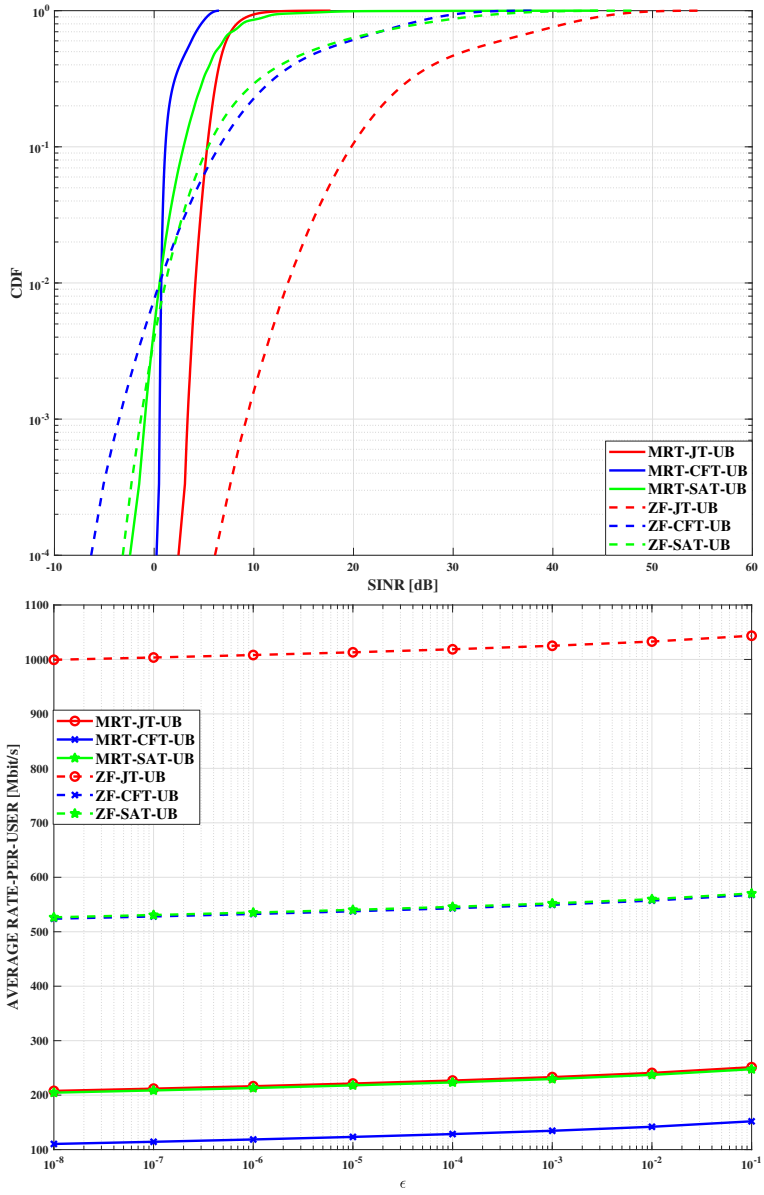



FIGURE 4.5: On the top we have the SINR CDF and on the bottom the average rate-per-user with MRT versus FZF for different transmission modes, JT, CFT, and SAT.

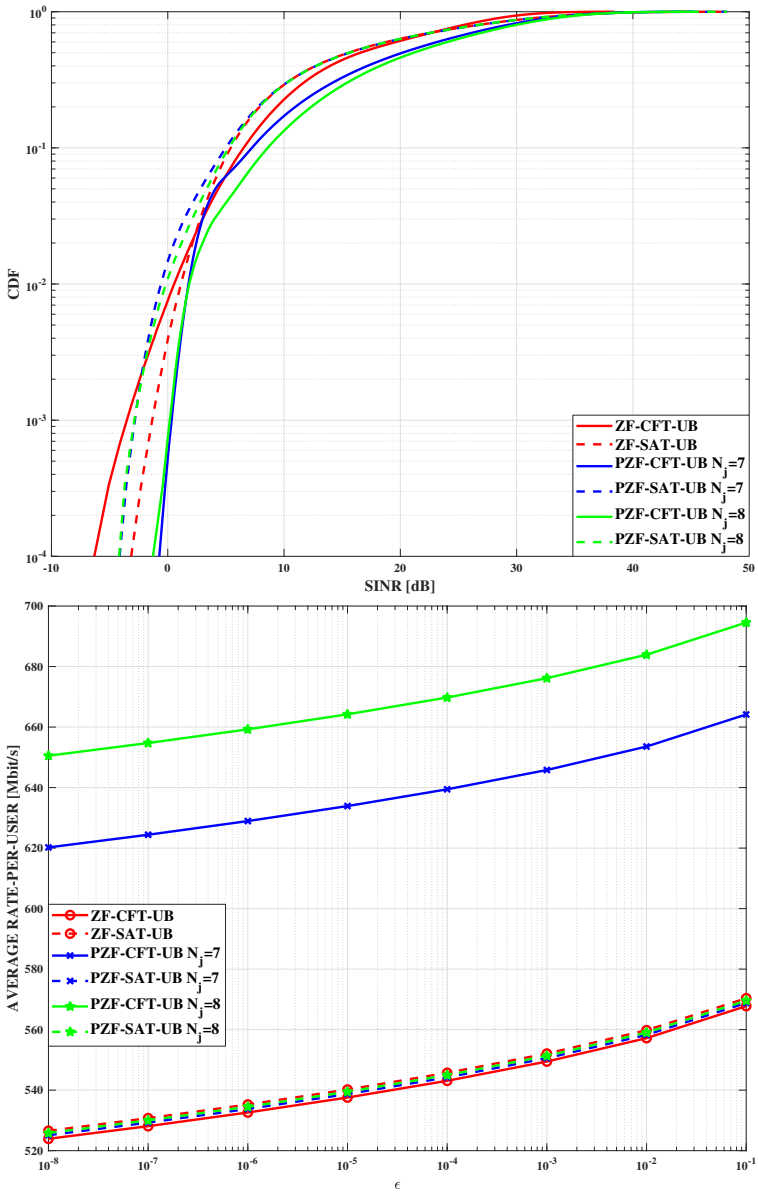



FIGURE 4.6: On the top we have the SINR CDF and on the bottom the average rate-per-user with FZF versus PZF with $N_j = 7$ and $N_j = 8$ for for different transmission modes, CFT and SAT.

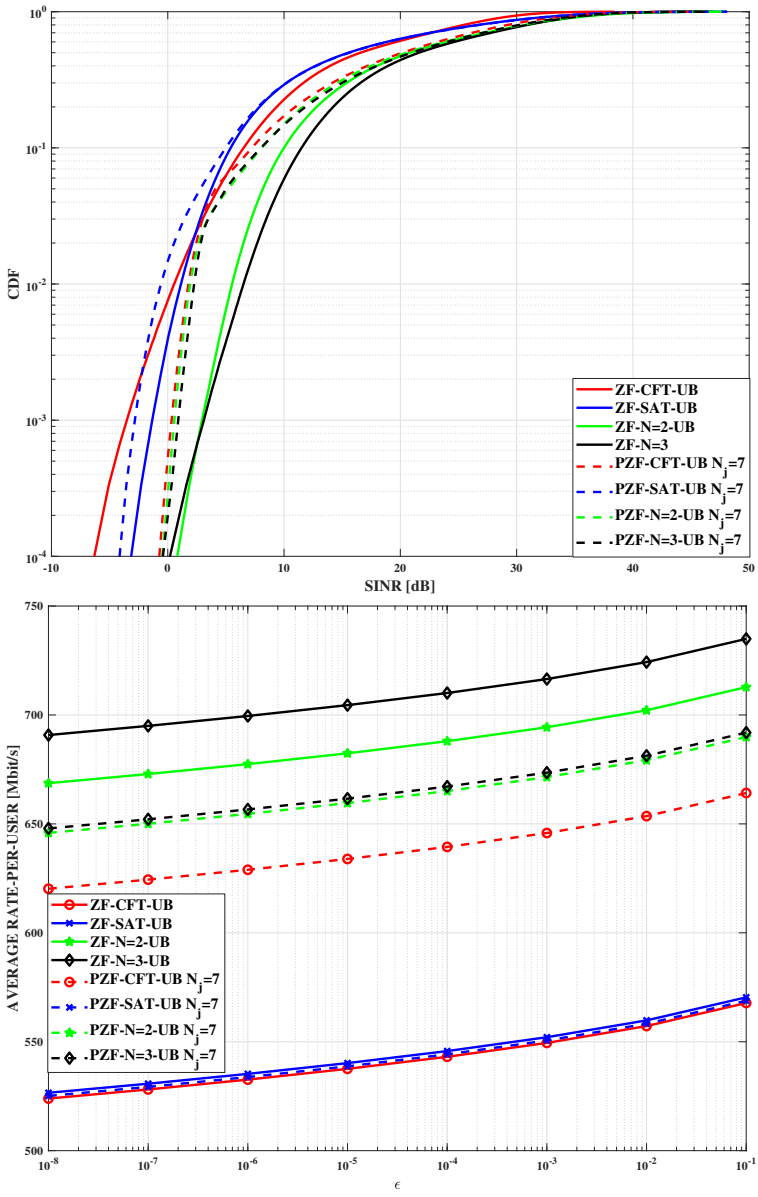



FIGURE 4.7: On the top we have the SINR CDF and on the bottom the average rate-per-user with FZF versus PZF with $N_j = 7$ for for different transmission modes, CFT, SAT, UC with N=2 and N=3.

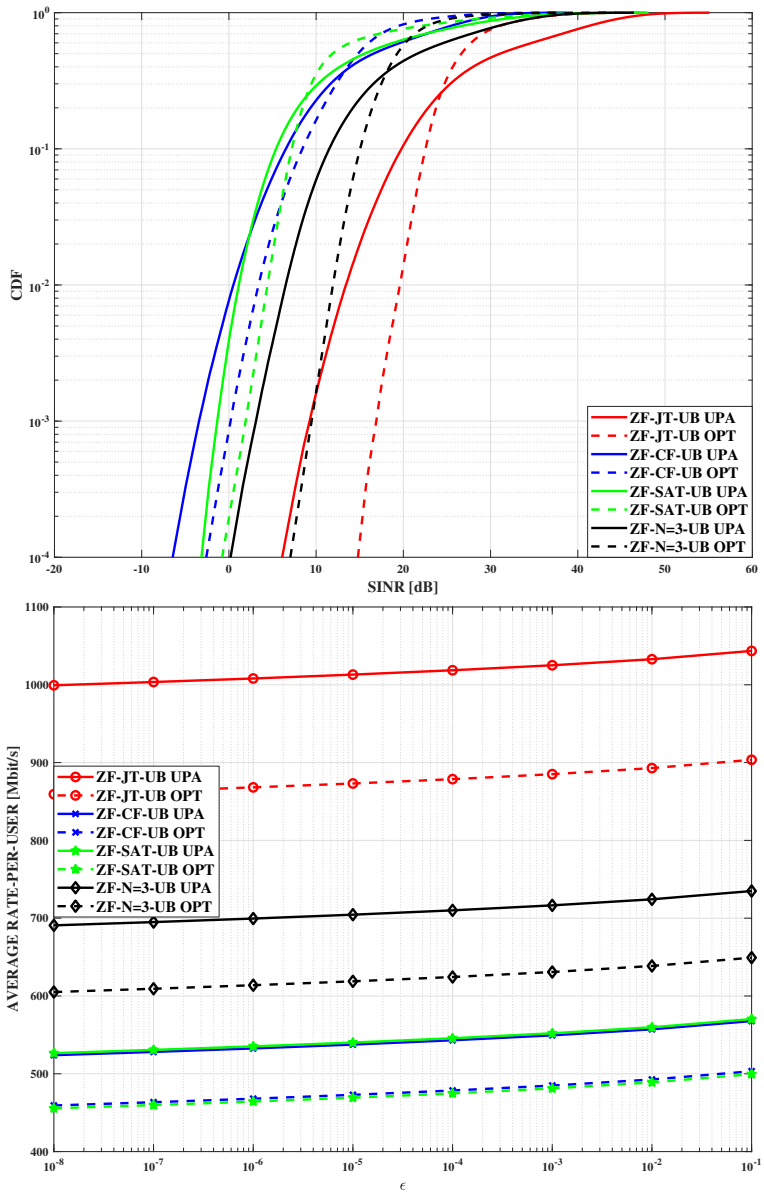



FIGURE 4.8: On the top we have the SINR CDF and on the bottom the average rate-per-user with FZF with equal and optimal power allocation, CFT, SAT, UC with N=3.

# Chapter 5

# Conclusions

In this thesis the potential of massive MIMO systems has been highlighted. Of course, this technology is continuously evolving, enhancing its performances. Massive MIMO systems, as written in the introduction, play a fondamental role in the 5G and beyond-5G wireless networks, since they permit to improve the communication's quality. In the next future, a lot of devices will be simultaneously connected, as in vehicular communications, in mMTC or in wireless sensor networks. Thus, there are different use-cases to be considered, and each of them has different requirements and quality of services. Here, the conclusions of the Ph.D thesis are reported, by reassuming what has been shown in the previous chapters.

In chapter 1 an overview of 5G wireless networks has been introduced by highlighting its features, use-cases and new technology to be exploited. Then the contributions of the thesis have been presented.

In chapter 2 the advantages given by the usage of massive MIMO system has been introduced. Moreover, in the same chapter the massive MIMO has even been analyzed from the mathematical point of view. Thus, all the phases of the signal processing required in these system have been reported.

In chapter 3 the CF massive MIMO system operating at mmWave frequencies with a UC association between APs and MSs has been considered. Adopting a clustered channel model capable of taking into account the channel correlation for nearby devices, the paper has analyzed both the CF and UC approaches, by proposing a low-complexity power allocation rule aimed at

global energy efficiency maximization. The presence of HY beamforming architectures at the APs has been considered, whereas, for the MSs, a simple channel-independent 0-1 beamforming structure has been considered. The obtained results have confirmed that the proposed resource allocation algorithms are effective at increasing the system energy efficiency and the system average rate-per-user, as well as that the use of HY beamforming architectures introduces, as expected, a considerable performance degradation.

Chapter 4 has considered the adoption of distributed antenna settings in an indoor industrial scenario. Several transmission modes and several beamformers have been adopted. A power control rule aimed at maximizing the minimum SINR across users has been considered, which is extremely useful in industrial settings where reliability is to be privileged with respect to peak data rates. Overall, results have shown that the user centric transmission mode, wherein each AC is served by a limited number of APs in the neighbors, achieves the best trade-off between implementation complexity and system performance.

Moreover the industrial scenarios could be also studied from a different point of view. In [7], the same industrial scenario has been considered but the analysis was different. It has been shown the potential of D-MIMO systems in terms of BLER by considering different beamformers and transmission modes. Moreover, as future developments, latency and BLER analysis can be carried out in conjunction with the power control.

# Appendix A

# Successive lower-bound optimization

In this appendix we briefly review the successive lower-bound optimization technique. The method is based on the idea of merging the tools of alternating optimization [60, Section 2.7] and sequential convex programming [61]. Specifically, let us consider the generic optimization problem

$$\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \,, \tag{A.1}$$

with $f : \mathbb{R}^n \to \mathbb{R}$ a differentiable function, and $\mathcal{X}$ a compact set. Similarly to the alternating optimization method, the successive lower-bound maximization partitions the variable space into $J$ blocks, $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J)$, which are cyclically optimized one at a time, while the other blocks are kept fixed. Thus, Problem (A.1) is decomposed into $M$ subproblems, wherein the generic subproblem is stated as

$$\max_{\boldsymbol{x}_m} f(\boldsymbol{x}_m, \boldsymbol{x}_{-m}) \,, \tag{A.2}$$

with $\boldsymbol{x}_{-m}$ cotaining all variable blocks except the $m$-th. If Problem (A.2) is globally solved for each $m = 1, \ldots, J$, then we have an instance of the alternating maximization method, which, as proved in [60, Proposition 2.7.1], monotonically improves the objective of (A.1), and yields a first-order optimal point if the solution of (A.2) is unique for any $m$, and if $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_J$,

with $\boldsymbol{x}_m \in \mathcal{X}_m$ for all $m$. However, if globally solving (A.2) is difficult (e.g. because (A.2) is not a convex problem), then implementing the alternating maximization method proves difficult. In this case, the successive lower-bound maximization method proposes to find a (possibly suboptimal) solution of (A.2), by means of the sequential convex programming method. Besides leading to a computationally viable algorithm, this approach is proved to preserve the optimality properties of the alternating optimization method [37], despite the fact that a possible suboptimal solution of (A.2) is determined.

As for sequential optimization, its basic idea is to tackle a difficult maximization problem by a sequence of easier maximization problems. To elaborate, denote by $g_i(\boldsymbol{x}_m)$ the $i$-th constraint of (A.2), for $i = 1, \ldots, C$, and 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 $\ell$:

**(P1)** $f_\ell(\boldsymbol{x}_m) \leq f(\boldsymbol{x}_m)$, $g_{i,\ell}(\boldsymbol{x}_m) \leq g_{i,\ell}(\boldsymbol{x}_m)$, 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 [61] (see also [37], [62]) it is shown that, subject to constraint qualifications, the sequence $\{f(\boldsymbol{x}_m^{(\ell)})\}_\ell$ of the solutions of the $\ell$-th Problem $\mathcal{P}_\ell$, is monotonically increasing and converges. Moreover, every convergent sequence $\{\boldsymbol{x}_m^{(\ell)}\}_\ell$ attains a first-order optimal point of the original Problem (A.2). Thus, sequential optimization guarantees at the same time to monotonically improve the objective function, and to fulfill the KKT first-order optimality conditions of the original problem. Nevertheless, in order to be able to use the method, it is necessary to find lower bounds of the original objective function, which fulfill all three properties **P1**, **P2**, **P3**, while at the same time leading to an approximate problem that can be solved with affordable complexity.

# Appendix B

# Derivation of the Lower Bound for UCT

We have to find a closed form for the terms $D_k$, $B_k$, and $I_{k,j}$.

**Evaluation of $\mathbf{D}_k$**

Denoting by $\tilde{\mathbf{h}}_{k,j} = \mathbf{h}_{k,j} - \hat{\mathbf{h}}_{k,j}$ the channel estimation error, and it is known, from the LMMSE estimation, that $\tilde{\mathbf{h}}_{k,j}$ $\hat{\mathbf{h}}_{k,j}$ are independent. So, by substituting $\mathbf{h}_{k,j} = \tilde{\mathbf{h}}_{k,j} + \hat{\mathbf{h}}_{k,j}$, it is possible to derive the following expression:

$$\begin{aligned} D_k = \mathbb{E}\left[\sum_{j\in J(k)} \sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H \hat{\mathbf{h}}_{k,j}\right] = \sum_{j\in J(k)} \sqrt{\eta_{k,j}}\mathbb{E}\left[\mathbf{h}_{k,j}^H \hat{\mathbf{h}}_{k,j}\right] = \\ \sum_{j\in J(k)} \sqrt{\eta_{k,j}}\left(\mathbb{E}\left[\hat{\mathbf{h}}_{k,j}^H \hat{\mathbf{h}}_{k,j}\right] + \mathbb{E}\left[\tilde{\mathbf{h}}_{k,j}^H \hat{\mathbf{h}}_{k,j}\right]\right) = \sum_{j\in J(k)} \sqrt{\eta_{k,j}}\gamma_{k,j} \end{aligned} \tag{B.1}$$

In eq.(B.1) $\gamma_{k,j}$ is defined as:

$$\begin{aligned} \gamma_{k,j} = \mathbb{E}\left[\hat{\mathbf{h}}_{k,j}^H \hat{\mathbf{h}}_{k,j}\right] = tr\left(\mathbb{E}\left[\hat{\mathbf{h}}_{k,j}\hat{\mathbf{h}}_{k,j}^H\right]\right) = tr\left(\mathbb{E}\left[\alpha_{k,j}\check{\mathbf{y}}_{k,j}\check{\mathbf{y}}_{k,j}^H\alpha_{k,j}\right]\right) = \\ = tr\Bigg(\alpha_{k,j}\underbrace{\mathbb{E}\left[\check{\mathbf{y}}_{k,j}\check{\mathbf{y}}_{k,j}^H\right]\mathbb{E}\left[\check{\mathbf{y}}_{k,j}\check{\mathbf{y}}_{k,j}^H\right]^{-1}}_{\mathbf{I}_M}\mathbb{E}\left[\check{\mathbf{y}}_{k,j}\mathbf{h}_{k,j}^H\right]\Bigg) = \sqrt{\eta_k^{UL-T}}\beta_{k,j}\alpha_{k,j}M. \end{aligned} \tag{B.2}$$

**Evaluation of** $\mathbb{E}\left[|B_k|^2\right]$

Since the variance of a sum of independent RVs is equal to the sum of their variances, $\mathbb{E}\left[|B_k|^2\right]$ can be written as:

$$\begin{aligned}
\mathbb{E}\left[|B_k|^2\right] &= \mathbb{E}\left[\left|\sum_{j\in J(k)}\sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j} - \mathbb{E}\left[\sum_{j\in J(k)}\sqrt{\eta_{k,j}}\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]\right|^2\right] = \\
&= \mathbb{E}\left[\left|\sum_{j\in J(k)}\sqrt{\eta_{k,j}}\left(\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j} - \mathbb{E}\left[\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]\right)\right|^2\right] = \\
&= \sum_{j\in J(k)}\eta_{k,j}\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j} - \mathbb{E}\left[\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]\right|^2\right] = \\
&= \sum_{j\in J(k)}\eta_{k,j}\left(\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right|^2\right] + \left|\mathbb{E}\left[\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]\right|^2 + \right. \\
&\left. - 2\mathrm{Re}\mathbb{E}\left[\left(\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right)\mathbb{E}\left[\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]\right]\right) = \\
&= \sum_{j\in J(k)}\eta_{k,j}\left(\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right|^2\right] - \left|\underbrace{\mathbb{E}\left[\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right]}_{\gamma_{k,j}}\right|^2\right)
\end{aligned} \tag{B.3}$$

Now we can evaluate the term $\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right|^2\right]$:

$$\begin{aligned}
&\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\hat{\mathbf{h}}_{k,j}\right|^2\right] = \mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\alpha_{k,j}\left(\sum_{l=1}^{K}\sqrt{\eta_l^{UL-T}}\mathbf{h}_{l,j}\boldsymbol{\Phi}_{:,l}^H\boldsymbol{\Phi}_{:,k}+\tilde{\mathbf{w}}_{k,j}\right)\right|^2\right] = \\
&= \mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\alpha_{k,j}\sqrt{\eta_k^{UL-T}}\mathbf{h}_{k,j}\right|^2\right] + \\
&+ \mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\alpha_{k,j}\sum_{l=1,l\neq k}^{K}\sqrt{\eta_l^{UL-T}}\mathbf{h}_{l,j}\boldsymbol{\Phi}_{:,l}^H\boldsymbol{\Phi}_{:,k}\right|^2\right] + \mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\alpha_{k,j}\tilde{\mathbf{w}}_{k,j}\right|^2\right] = \\
&= \alpha_{k,j}^2\eta_k^{UL-T}\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right] + \sum_{l=1,l\neq k}^{K}\eta_l^{UL-T}\alpha_{k,j}^2\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{l,j}\right|^2\right]\left|\boldsymbol{\Phi}_{:,l}^H\boldsymbol{\Phi}_{:,k}\right|^2 + \\
&+ \alpha_{k,j}^2\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\tilde{\mathbf{w}}_{k,j}\right|^2\right] = \alpha_{k,j}^2\eta_k^{UL-T}\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right] + \\
&+ \sum_{l=1,l\neq k}^{K}\eta_l^{UL-T}\alpha_{k,j}^2\mathbb{E}\left[\mathbf{h}_{k,j}^H\mathbf{h}_{l,j}\mathbf{h}_{l,j}^H\mathbf{h}_{k,j}\right]\left|\boldsymbol{\Phi}_{:,l}^H\boldsymbol{\Phi}_{:,k}\right|^2 + \alpha_{k,j}^2\mathbb{E}\left[\mathbf{h}_{k,j}^H\tilde{\mathbf{w}}_{k,j}\tilde{\mathbf{w}}_{k,j}^H\mathbf{h}_{k,j}\right] = \\
&= \alpha_{k,j}^2\eta_k^{UL-T}\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right] + \\
&+ \sum_{l=1,l\neq k}^{K}\eta_l^{UL-T}\alpha_{k,j}^2\beta_{l,j}\beta_{k,j}M\left|\boldsymbol{\Phi}_{:,l}^H\boldsymbol{\Phi}_{:,k}\right|^2 + \alpha_{k,j}^2\beta_{k,j}\sigma_w^2M
\end{aligned} \tag{B.4}$$

where $\tilde{\mathbf{w}}_{k,j} = \mathbf{W}_j \mathbf{\Phi}_{:,k}$. We have to evaluate the terms $\mathbb{E}\left[\mathbf{h}_{k,j}^H \mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j}\right]$ and $\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H \mathbf{h}_{k,j}\right|\right]$, and they are equal to:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{h}_{k,j}^H \mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j}\right] &= tr\left(\mathbb{E}\left[\mathbf{h}_{k,j}^H \mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j}\right]\right) = \mathbb{E}\left[tr\left(\mathbf{h}_{k,j}^H \mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j}\right)\right] = \\
&= \mathbb{E}\left[tr\left(\mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j} \mathbf{h}_{k,j}^H\right)\right] = \\
&= tr\left(\mathbb{E}\left[\mathbf{h}_{l,j} \mathbf{h}_{l,j}^H \mathbf{h}_{k,j} \mathbf{h}_{k,j}^H\right]\right) = \\
&= tr\Big(\underbrace{\mathbb{E}\left[\mathbf{h}_{l,j} \mathbf{h}_{l,j}^H\right]}_{\beta_{l,j}\mathbf{I}_M} \underbrace{\mathbb{E}\left[\mathbf{h}_{k,j} \mathbf{h}_{k,j}^H\right]}_{\beta_{k,j}\mathbf{I}_M}\Big) = \beta_{k,j}\beta_{l,j} M
\end{aligned} \tag{B.5}$$

and

$$\begin{aligned}
\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H \mathbf{h}_{k,j}\right|\right] &= \mathrm{tr}^2(\beta_{k,j}\mathbf{I}_M) + \mathrm{tr}(\beta_{k,j}\mathbf{I}_M \beta_{k,j}\mathbf{I}_M) = \\
&= \beta_{k,j}^2 M^2 + \beta_{k,j}^2 M = \beta_{k,j}^2 (M+1) M
\end{aligned} \tag{B.6}$$

Substituting Eq. (B.5) and (B.6) in Eq. (B.4) first, and then in Eq. (B.3) we get:

$$\mathbb{E}\left[|B_k|^2\right] = \sum_{j \in J(k)} \eta_{k,j} \left(\gamma_{k,j}^2 + \gamma_{k,j}\beta_{k,j} - \gamma_{k,j}^2\right) = \sum_{j \in J(k)} \eta_{k,j}\gamma_{k,j}\beta_{k,j} \tag{B.7}$$

**Evaluation of** $\mathbb{E}\left[|I_{k,l}|^2\right]$

As it has been done for $\mathbb{E}\left[|B_k|^2\right]$, it is possible to follows a similar approach.

$$
\begin{aligned}
&\mathbb{E}\left[|I_{k,l}|^2\right] = \mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\mathbf{h}_{k,j}^H \hat{\mathbf{h}}_{l,j}\right|^2\right] = \\
&= \mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}^H \left(\sum_{i=1}^{K} \eta_i^{UL-T}\mathbf{h}_{i,j}\boldsymbol{\Phi}_{:,i}^H\boldsymbol{\Phi}_{:,l} + \tilde{\mathbf{W}}_{l,j}\right)\right|^2\right] = \\
&= \mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}^H \eta_k^{UL-T}\mathbf{h}_{k,j}\boldsymbol{\Phi}_{:,i}^H\boldsymbol{\Phi}_{:,l}\right|^2\right] + \\
&+ \mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}^H \left(\sum_{i=1,i\neq k}^{K} \sqrt{\eta}_i^{UL-T}\mathbf{h}_{i,j}\boldsymbol{\Phi}_{:,i}^H\boldsymbol{\Phi}_{:,l}\right)\right|^2\right] + \\
&+ \mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}^H \tilde{\mathbf{W}}_{l,j}\right|^2\right] = \eta_k^{UL-T}\mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}\mathbf{h}_{k,j}\right|^2\right]\left|\boldsymbol{\Phi}_{:,k}^H\boldsymbol{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{j\in J(l)}\sum_{i=1,i\neq k}^{K} \eta_{l,j}\eta_i^{UL-T}\alpha_{l,j}^2\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{i,j}\right|^2\right]\left|\boldsymbol{\Phi}_{:,i}^H\boldsymbol{\Phi}_{:,l}\right|^2 + \sum_{j\in J(l)} \eta_{l,j}\alpha_{l,j}^2\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\tilde{\mathbf{W}}_{l,j}\right|^2\right] = \\
&= \eta_k^{UL-T}\mathbb{E}\left[\left|\sum_{j\in J(l)} \sqrt{\eta_{l,j}}\alpha_{l,k}\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right]\left|\boldsymbol{\Phi}_{:,k}^H\boldsymbol{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{j\in J(l)}\sum_{i=1,i\neq k}^{K} \eta_{l,j}\eta_i^{UL-T}\alpha_{l,j}^2\beta_{i,j}\beta_{k,j}M\left|\boldsymbol{\Phi}_{:,k}^H\boldsymbol{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{j\in J(l)} \eta_{l,j}\alpha_{l,j}^2\beta_{k,j}\sigma_w^2 M
\end{aligned}
\tag{B.8}
$$

At this point, $\mathbb{E}\left[\left|\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,k}\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right]$ has to be evaluated:

$$
\begin{aligned}
&\mathbb{E}\left[\left|\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right] = \\
&= \mathbb{E}\left[\left(\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}^2\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right)\left(\sum_{j'\in J(l)}\sqrt{\eta_{l,j'}}\alpha_{l,j'}\mathbf{h}_{k,j'}^H\mathbf{h}_{k,j'}\right)\right] = \\
&= \sum_{j\in J(l)}\eta_{l,j}\alpha_{l,j}^2\mathbb{E}\left[\left|\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\right|^2\right] + \\
&+ \sum_{j\in J(l)}\sum_{j'\in J(l),j'\neq j}\sqrt{\eta_{l,j}}\sqrt{\eta_{l,j'}}\alpha_{l,j}\alpha_{l,j'}\mathbb{E}\left[\mathbf{h}_{k,j}^H\mathbf{h}_{k,j}\mathbf{h}_{k,j'}^H\mathbf{h}_{k,j'}^H\right] = \\
&= \sum_{j\in J(l)}\eta_{l,j}\alpha_{l,j}^2\left(\beta_{k,j}^2M^2+\beta_{k,j}^2M\right) + \\
&+ \sum_{j\in J(l)}\sum_{j'\in J(l),j'\neq j}\sqrt{\eta_{l,j}}\sqrt{\eta_{l,j'}}\alpha_{l,j}\alpha_{l,j'}\beta_{k,j}\beta_{k,j'}M^2 = \\
&= \sum_{j\in J(l)}\eta_{l,j}\alpha_{l,j}^2\beta_{k,j}^2M + \left(\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}\beta_{k,j}M\right)^2
\end{aligned}
\tag{B.9}
$$

Substituting, now, Eq. (B.9) in Eq. (B.8) we get:

$$\begin{aligned}
&\mathbb{E}\left[|I_{k,l}|^2\right] = \eta_k^{UL-T}\left[\sum_{j\in J(l)}\eta_{l,j}\alpha_{l,j}^2\beta_{k,j}^2 M + \left(\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}\beta_{k,j}M\right)^2\right]\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{j\in J(l)}\sum_{i=1,i\neq k}^{K}\eta_{l,j}\eta_i^{UL-T}\alpha_{l,j}^2\beta_{i,j}\beta_{k,j}M\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right|^2 + \sum_{j\in J(l)}\eta_{l,j}\alpha_{l,j}^2\beta_{k,j}\sigma_w^2 M = \\
&= \sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}^2\beta_{k,j}M\underbrace{\left(\sum_{i=1}^{K}\eta_i^{UL-T}\beta_{i,j}\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right| + \sigma_w^2\right)}_{\mathrm{den}(\alpha_{l,j})} + \\
&+ \left(\sum_{j\in J(l)}\sqrt{\eta_k^{UL-T}}\sqrt{\eta_{l,j}}\alpha_{l,j}\beta_{k,j}M\right)^2\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 = \\
&= \sum_{j\in J(l)}\eta_{l,j}\beta_{k,j}\frac{\overbrace{\sqrt{\eta_l^{UL-T}}\alpha_{l,j}\beta_{l,j}M}^{\gamma_{l,j}}}{\mathrm{den}(\alpha_{l,j})}\mathrm{den}(\alpha_{l,j}) + \left(\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\alpha_{l,j}\beta_{k,j}M\right)^2\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 = \\
&= \sum_{j\in J(l)}\eta_{l,j}\beta_{k,j}\alpha_{l,j} + \left(\sum_{j\in J(l)}\sqrt{\eta_{l,j}}\frac{\sqrt{\eta_k^{UL-T}}}{\sqrt{\eta_k^{UL-T}}}\frac{\beta_{k,j}}{\beta_{l,j}}\alpha_{l,j}\right)^2\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2
\end{aligned} \tag{B.10}$$

where $\mathrm{den}(\alpha_{k,j})$ is the denominator of $\alpha_{k,j}$ defined in Eq. (4.5). Finally, we can plug Eq. (B.1), Eq. (B.7) and Eq.(B.10) into Eq. (4.17) to obtain Eq. (4.18).

# Appendix C

# Derivation of the Lower Bound for Distributed MIMO

**Definition of block matrices**

It is possible to write $\hat{\mathbf{h}}_k = \mathbf{D}_k \check{\mathbf{y}}_k$, where $\mathbf{D}_k$ is a $(M_{TOT} \times M_{TOT})$ block diagonal matrix defined in Section 4.4 and $\check{\mathbf{y}}_k$ is defined as:

$$\check{\mathbf{y}}_k = \begin{bmatrix} \check{\mathbf{y}}_{k,1} \\ \check{\mathbf{y}}_{k,2} \\ \vdots \\ \check{\mathbf{y}}_{k,J} \end{bmatrix} \tag{C.1}$$

So we have the following power constraint: $\eta_k \mathbb{E}\left[\hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_k\right] \leq P_{max}$. We know that $\mathbb{E}\left[\hat{\mathbf{h}}_k^H \hat{\mathbf{h}}_k\right] = \mathrm{tr}\left(\mathbb{E}\left[\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H\right]\right)$ so we can write $\mathbb{E}\left[\hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H\right] = \mathbb{E}\left[\mathbf{D}_k \check{\mathbf{y}}_k \check{\mathbf{y}}_k^H \mathbf{D}_k^H\right]$.

**Evaluation of $\mathbb{E}\left[\check{\mathbf{y}}_k \check{\mathbf{y}}_k^H\right]$**

It is possible to evaluate separately the quantity $\mathbb{E}\left[\check{\mathbf{y}}_k \check{\mathbf{y}}_k^H\right]$, so first of all we define:

$$\check{\mathbf{y}}_{k,j} = \sqrt{\eta_k^{UL-T}} \mathbf{h}_{k,j} + \sum_{l \neq k} \sqrt{\eta_l^{UL-T}} \mathbf{h}_{l,j} \boldsymbol{\Phi}_{:,l}^H \boldsymbol{\Phi}_{:,k} + \tilde{\mathbf{W}}_{k,j}, \tag{C.2}$$

then,

$$\mathbb{E}\left[\check{\mathbf{y}}_{k,j}\check{\mathbf{y}}_{k,j}^H\right] = \left(\sum_{l=1}^{K}\eta_l^{UL-T}\beta_{l,j}\left|\mathbf{\Phi}_{:,l}^H\mathbf{\Phi}_{:,k}\right|^2 + \sigma_w^2\right)\mathbf{I}_M = c_{k,j}\mathbf{I}_M. \tag{C.3}$$

Now, assuming that the channels among users and APs are independent:

$$\mathbb{E}\left[\check{\mathbf{y}}_k\check{\mathbf{y}}_k^H\right] = \mathbf{C}_k = \begin{bmatrix} c_{k,1}\mathbf{I}_M & 0 & \dots & 0 \\ 0 & c_{k,2}\mathbf{I}_M & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_{k,J}\mathbf{I}_M \end{bmatrix} \tag{C.4}$$

Since $\alpha_{k,j} = \frac{\sqrt{\eta_k^{UL-T}}\beta_{k,j}}{c_{k,j}}$ and we are dealing with diagonal matrices we can rewrite, $\mathbf{D}_k = \sqrt{\eta_k^{UL-T}}\mathbf{B}_k\mathbf{C}_k^{-1}$, where $\mathbf{B}_k$ is defined in Section 4.4.

In the end, $\sqrt{\eta_k^{UL-T}}\hat{\mathbf{h}}_k = \mathbf{B}_k\mathbf{C}_k^{-1}\check{\mathbf{y}}_k$ and so:

$$\mathbb{E}\left[\hat{\mathbf{h}}_k^H\hat{\mathbf{h}}_k\right] = \mathbf{B}_k\mathbf{C}_k^{-1}\mathbb{E}\left[\check{\mathbf{y}}_{k,j}\check{\mathbf{y}}_{k,j}^H\right]\mathbf{C}_k^{-1}\mathbf{B}_k^H = \sqrt{\eta_k^{UL-T}}\mathbf{D}_k\mathbf{B}_k^H. \tag{C.5}$$

**Evaluation of $D_k$**

By looking at the eq.(4.21) $D_k$ can be written as:

$$\begin{aligned} D_k &= \mathbb{E}\left[\mathbf{h}_k^H\hat{\mathbf{q}}_k\right] = \mathbb{E}\left[\mathbf{h}_k^H\hat{\mathbf{h}}_k\right] = \mathbb{E}\left[\mathbf{h}_k^H\mathbf{D}_k\check{\mathbf{y}}_k\right] = \\ &= \mathrm{tr}\left(\mathbb{E}\left[\check{\mathbf{y}}_k\mathbf{h}_k^H\right]\mathbf{D}_k\right) = \mathrm{tr}\left(\mathbb{E}\left[\left(\sum_{l=1}^{K}\sqrt{\eta_l^{UL-T}}\mathbf{h}_{l,j}\mathbf{\Phi}_{:,l}^H\mathbf{\Phi}_{:,k} + \tilde{\mathbf{W}}_{k,j}\right)\mathbf{h}_k^H\right]\mathbf{D}_k\right) = \\ &= \mathrm{tr}\left(\sqrt{\eta_k^{UL-T}}\mathbb{E}\left[\mathbf{h}_k\mathbf{h}_k^H\right]\mathbf{D}_k\right) = \gamma_k \end{aligned} \tag{C.6}$$

**Evaluation of** $\mathbb{E}\left[|B_k|^2\right]$

Now, we have to evaluate $\mathbb{E}\left[|B_k|^2\right]$.

$$\mathbb{E}\left[|B_k|^2\right] = \mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{q}_k - \mathbb{E}\left[\mathbf{h}_k^H\mathbf{q}_k\right]\right|^2\right] = \mathbb{E}\left[\left|\mathbf{h}_k^H\hat{\mathbf{h}}_k\right|^2\right] - \left|\underbrace{\mathbb{E}\left[\mathbf{h}_k^H\hat{\mathbf{h}}_k\right]}_{\gamma_k}\right|^2 \tag{C.7}$$

Assuming that all the channels are independent it is possible to follow the same approach for the UC; then we can write:

$$\mathbb{E}\left[|B_k|^2\right] = \sqrt{\eta_k^{UL-T}}\mathrm{tr}(\mathbf{B}_k\mathbf{D}_k^H\mathbf{B}_k). \tag{C.8}$$

**Evaluation of** $\mathbb{E}\left[|I_{k,l}|^2\right]$

From eq.(4.21) we have to evaluate $\mathbb{E}\left[\left|\mathbf{h}_k^H\hat{\mathbf{h}}_l\right|^2\right]$:

$$
\begin{aligned}
\mathbb{E}\left[|I_{k,l}|^2\right] &= \mathbb{E}\left[\left|\mathbf{h}_k^H\hat{\mathbf{h}}_l\right|^2\right] = \mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{D}_l\check{\mathbf{y}}_l\right|^2\right] = \\
&= \mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{D}_l(\sum_{i=1}^{K}\sqrt{\eta_i^{UL-T}}\mathbf{h}_{i,j}\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,k}+\tilde{\mathbf{W}}_l)\right|^2\right] = \\
&= \eta_k^{UL-T}\mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{D}_l\mathbf{h}_k\right|^2\right]\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{i=1,i\neq k}^{K}\eta_i^{UL-T}\mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{D}_l\mathbf{h}_i\right|^2\right]\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right|^2 + \mathbb{E}\left[\left|\mathbf{h}_k^H\mathbf{D}_l\tilde{\mathbf{W}}_l\right|^2\right] = \\
&= \eta_k^{UL-T}\left[\mathrm{tr}^2\left(\mathbf{D}_l\mathbf{B}_k\right)+\mathrm{tr}\left(\mathbf{D}_l\mathbf{B}_k\mathbf{D}_l^H\mathbf{B}_k\right)\right]\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \\
&+ \sum_{i=1,i\neq k}^{K}\eta_i^{UL-T}\mathrm{tr}\left(\mathbf{D}_l\mathbf{B}_i\mathbf{D}_l^H\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right|^2 + \sigma_w^2\mathrm{tr}\left(\mathbf{D}_l\mathbf{D}_l^H\mathbf{B}_k\right) = \\
&= \eta_k^{UL-T}\mathrm{tr}^2\left(\mathbf{D}_l\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \sum_{i=1}^{K}\eta_i^{UL-T}\mathrm{tr}\left(\mathbf{D}_l\mathbf{B}_i\mathbf{D}_l^H\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right|^2 + \\
&+ \sigma_w^2\mathrm{tr}\left(\mathbf{D}_l\mathbf{D}_l^H\mathbf{B}_k\right) \overset{(a)}{=} \\
&\overset{(a)}{=} \eta_k^{UL-T}\mathrm{tr}^2\left(\mathbf{D}_l\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \\
&+ \mathrm{tr}\left[\mathbf{D}_l\Big(\underbrace{\sum_{i=1}^{K}\eta_i^{UL-T}\mathbf{B}_i\left|\mathbf{\Phi}_{:,i}^H\mathbf{\Phi}_{:,l}\right|^2+\sigma_w^2\mathbf{I}_{M_{TOT}}}_{\mathbf{C}_l}\Big)\mathbf{D}_l^H\mathbf{B}_k\right] = \\
&= \eta_k^{UL-T}\mathrm{tr}^2\left(\mathbf{D}_l\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \mathrm{tr}\left(\mathbf{D}_l\mathbf{C}_l\mathbf{D}_l^H\mathbf{B}_k\right) \overset{(b)}{=} \\
&\overset{(b)}{=} \eta_k^{UL-T}\mathrm{tr}^2\left(\mathbf{D}_l\mathbf{B}_k\right)\left|\mathbf{\Phi}_{:,k}^H\mathbf{\Phi}_{:,l}\right|^2 + \sqrt{\eta_l^{UL-T}}\mathrm{tr}\left(\mathbf{B}_l\mathbf{D}_l^H\mathbf{B}_k\right)
\end{aligned}
\tag{C.9}
$$

where in (a) we exploited the linearity of the trace and in (b) we exploited $\mathbf{D}_l = \sqrt{\eta_l^{UL-T}}\mathbf{B}_l\mathbf{C}_l^{-1}$. In the end, summing up all the terms in the eq.(C.6),(C.8) and (C.9) we get eq.(4.23).

# Bibliography

[1] A. D. Zayas, A. F. Cattoni, C. Cardenas, O. Castañeda, J. Baños, P. Merino, and M. Dieudonne, "Triangle: A platform to validate 5g kpis in end to end scenarios", in *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, IEEE, 2018, pp. 1–7.

[2] H. Tullberg, P. Popovski, Z. Li, M. A. Uusitalo, A. Hoglund, O. Bulakci, M. Fallgren, and J. F. Monserrat, "The metis 5g system concept: Meeting the 5g requirements", *IEEE Communications magazine*, vol. 54, no. 12, pp. 132–139, 2016.

[3] S. Mattisson, "An overview of 5g requirements and future wireless networks: Accommodating scaling technology", *IEEE Solid-State Circuits Magazine*, vol. 10, no. 3, pp. 54–60, 2018. DOI: 10.1109/MSSC.2018.2844606.

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

[5] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-efficient power control in cell-free and user-centric massive mimo at millimeter wave", *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 3, pp. 651–663, 2019.

[6] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime", *IEEE Transactions on Information Theory*, vol. 56, no. 5, p. 2307, 2010.

[7] M. Alonzo, P. Baracca, S. R. Khosravirad, and S. Buzzi, "URLLC for factory automation: An extensive throughput-reliability analysis of D-MIMO", in *WSA 2020; 24th International ITG Workshop on Smart Antennas*, 2020, pp. 1–6.

[8] M. Alonzo, S. Buzzi, and A. Zappone, "Energy-efficient downlink power control in mmWave cell-free and user-centric massive MIMO", in *2018 IEEE 5G World Forum*, IEEE, 2018.

[9] M. Alonzo and S. Buzzi, "Cell-free and user-centric massive MIMO at millimeter wave frequencies", in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–5. DOI: 10.1109/PIMRC.2017.8292302.

[10] J. H. E. Björnson 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.

[11] D. Ring, "Mobile telephony-wide area coverage", *Bell Technical Laboratories technical memoranda*, 1947.

[12] 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, 2014.

[13] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone", in *2015 IEEE 16th international workshop on signal processing advances in wireless communications (SPAWC)*, IEEE, 2015, pp. 201–205.

[14] 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, 2017.

[15] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results", *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[16] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced", *IEEE Wireless Communications*, vol. 17, no. 3, 2010.

[17] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems", in *Proc. Asilomar Conference on Signals, Systems and Computers*, Citeseer, 2007, pp. 83–87.

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

[19] H. Q. Ngo, L. N. 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, 2018. DOI: 10.1109/TGCN.2017.2770215.

[20] S. Buzzi and C. D'Andrea, "User-centric communications versus cell-free massive MIMO for 5G cellular networks", in *Proc. 21st International ITG Workshop on Smart Antennas*, 2017.

[21] S. Buzzi and A. Zappone, "Downlink power control in user-centric and cell-free massive MIMO wireless networks", in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2017, pp. 1–6. DOI: 10.1109/PIMRC.2017.8292293.

[22] 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", *IEEE Transactions on Wireless Communications, submitted*, 2018.

[23] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?", pp. 1–7, 2016.

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

[25] 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!", vol. 1, pp. 335–349, May 2013.

[26] 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, 2018.

[27] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems", *IEEE Communications Magazine*, vol. 52, no. 12, pp. 122–131, 2014.

[28] "NGMN alliance 5G white paper", *https://www.ngmn.org/5g-white-paper/5g-white-paper.html*, 2015.

[29] 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. 5, 2016.

[30] S. Yun, S. H. Jeon, J. K. Choi, and A.-S. Park, "Energy efficiency of relay operation in millimeter-wave mobile broadband systems", in *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*, IEEE, 2014, pp. 1–0.

[31] 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, ISSN: 1567-2190. DOI: 10.1561/0100000088.

[32] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies", *IEEE Transactions on Signal Processing, in press*, vol. 64, no. 7, pp. 1668–1683, 2016.

[33] A. Pizzo and L. Sanguinetti, "Optimal design of energy-efficient millimeter wave hybrid transceivers for wireless backhaul", in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017 15th International Symposium on*, IEEE, 2017, pp. 1–8.

[34] S. Buzzi and C. D'Andrea, "On clustered statistical MIMO millimeter wave channel simulation", *ArXiv e-prints [Online] Available: https://arxiv.org/abs/1604.00648*, May 2016.

[35] A. Hjørungnes, *Complex-valued matrix derivatives: with applications in signal processing and communications*. Cambridge University Press, 2011.

[36] M. Iwanow, N. Vucic, M. H. Castaneda, J. Luo, W. Xu, and W. Utschick, "Some aspects on hybrid wideband transceiver design for mmwave communication systems", in *Smart Antennas (WSA 2016); Proceedings of the 20th International ITG Workshop on*, VDE, 2016, pp. 1–8.

[37] 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, pp. 1126–1153, 2013.

[38] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint cell selection and radio resource allocation in MIMO small cell networks via successive convex approximation", in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, IEEE, 2014, pp. 850–854.

[39] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization", *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4038–4050, 2013.

[40] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ Press, 2004.

[41] K. Haneda *et al.*, "5G 3GPP-like channel models for outdoor urban microcellular and macrocellular environments", *arXiv preprint. arXiv:1602.07533*, 2016.

[42] "5G Channel Model for bands up to 100 GHz", 2015. [Online]. Available: http://www.5gworkshops.com/5GCM.html.

[43] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the internet of things and Industry 4.0", *IEEE Industrial Electronics Magazine*, vol. 11, no. 1, pp. 17–27, 2017, ISSN: 1932-4529. DOI: 10.1109/MIE.2017.2649104.

[44] B. Holfeld, D. Wieruch, T. Wirth, L. Thiele, S. A. Ashraf, J. Huschke, I. Aktas, and J. Ansari, "Wireless communication for factory automation: An opportunity for LTE and 5G systems", *IEEE Communications Magazine*, vol. 54, no. 6, pp. 36–43, 2016, ISSN: 0163-6804. DOI: 10.1109/MCOM.2016.7497764.

[45] M. Gidlund, T. Lennvall, and J. Åkerberg, "Will 5g become yet another wireless technology for industrial automation?", in *2017 IEEE International Conference on Industrial Technology (ICIT)*, 2017, pp. 1319–1324. DOI: 10.1109/ICIT.2017.7915554.

[46] Č Stefanović, "Industry 4.0 from 5g perspective: Use-cases, requirements, challenges and approaches", in *2018 11th CMI International Conference: Prospects and Challenges Towards Developing a Digital Economy within the EU*, 2018, pp. 44–48. DOI: 10.1109/PCTDDE.2018.8624728.

[47] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale", *Proceedings of the IEEE*, vol. 106, no. 10, pp. 1834–1853, 2018, ISSN: 1558-2256. DOI: 10.1109/JPROC.2018.2867029.

[48] H. Chen, R. Abbas, P. Cheng, M. Shirvanimoghaddam, W. Hardjawana, W. Bao, Y. Li, and B. Vucetic, "Ultra-reliable low latency cellular networks: Use cases, challenges and approaches", *IEEE Communications*

*Magazine*, vol. 56, no. 12, pp. 119–125, 2018, ISSN: 1558-1896. DOI: 10.1109/MCOM.2018.1701178.

[49] 3rd Generation Partnership Project (3GPP), "Service requirements for the 5G system", Tech. Rep., TS 22.261, Sep. 2018.

[50] O. N. C. Yilmaz, Y. P. E. Wang, N. A. Johansson, N. Brahmi, S. A. Ashraf, and J. Sachs, "Analysis of ultra-reliable and low-latency 5G communication for a factory automation use case", in *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 1190–1195. DOI: 10.1109/ICCW.2015.7247339.

[51] A. E. Kalør, R. Guillaume, J. J. Nielsen, A. Mueller, and P. Popovski, "Network slicing for ultra-reliable low latency communication in industry 4.0 scenarios", *arXiv preprint arXiv:1708.09132*, 2017.

[52] X. Jiang, Z. Pang, M. Zhan, D. Dzung, M. Luvisotto, and C. Fischione, "Packet detection by a single OFDM symbol in URLLC for critical industrial control: A realistic study", *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 4, pp. 933–946, 2019, ISSN: 1558-0008. DOI: 10.1109/JSAC.2019.2898761.

[53] H. Ren, C. Pan, Y. Deng, M. Elkashlan, and A. Nallanathan, "Joint power and blocklength optimization for urllc in a factory automation scenario", *IEEE Transactions on Wireless Communications*, vol. 19, no. 3, pp. 1786–1801, 2020.

[54] G. Casciano, P. Baracca, and S. Buzzi, "Enabling ultra reliable wireless communications for factory automation with distributed MIMO", in *2019 IEEE Vehicular Technology Conference (VTC-Fall)*, 2019, pp. 1–7.

[55] R1-1813177, "Scenarios, frequencies and new field measurement results from two operational factory halls at 3.5 GHz for various antenna configurations", Nokia, Tech. Rep., Nov. 2018.

[56] W. Yang and G. Xu, "Optimal downlink power assignment for smart antenna systems", in *1998 IEEE International Conference on Acoustics, Speech*

*and Signal Processing (ICASSP)*, 1998, pp. 3337–3340. DOI: 10.1109/ICASSP.1998.679579.

[57] 3rd Generation Partnership Project (3GPP), "Study on channel model for frequencies from 0.5 to 100 GHz", Tech. Rep., TR 38.901, Jun. 2018.

[58] S. Verdu, *Multiuser detection*. Cambridge university press, 1998.

[59] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016. DOI: 10.1017/CBO9781316799895.

[60] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[61] 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, 1978.

[62] 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, 2010.

# *Acknowledgements*

On this last page of my PhD thesis I want to thank all the people who have been close to me in these three years of PhD. First of all, First of all I have to thank my supervisor, Prof. Stefano Buzzi, for his support from human and working point of view, because without him I would not have been able to reach this important goal. Then I want to thank Prof. Alessio Zappone, who in many occasions has helped me in carrying out my work. I have to thank Dr. Paolo Baracca, who followed me during my internship at the Nokia Bell Labs in Stuttgart. He made me feel so welcome into the work team and this allowed me to grow professionally and temperamentally. Then I thank all the others such as parents, sisters, friends, etc. for their support and in particular my fiancée, Maria Antoinetta who has been very close to me throughout this year.

# *Ringraziamenti (in italian)*

In quest'ultima pagina della mia tesi di dottorato voglio ringraziare tutte le persone che mi sono state vicine in questi tre anni di dottorato. Prima di tutti devo ringraziare il mio tutor, il Prof. Stefano Buzzi, per il suo sostegno sia dal punto di vista umano e sia dal punto di vista lavorativo perché senza di lui non avrei potuto raggiungere questo traguardo così importante. Poi voglio ringraziare il Prof.Alessio Zappone, anche lui in molte occasioni mi ha aiutato nello svolgimento del mio lavoro. Ringrazio, poi, il tutor che mi ha seguito durante la mia internship alla sede della Nokia di Stoccarda, il Dott. Paolo Baracca, che mi ha fatto subito integrare all'interno del team di lavoro e questo mi ha permesso di crescere sia professionalmente e sia caratterialmente. Ringrazio poi tutti gli altri, genitori, sorelle, amici, ecc. per il loro sostegno e in particolare la mia fidanzata, Maria Antonietta che mi è stata molto vicina durante tutto questo anno.