Contents

Notation

a	Scalar.
$\boldsymbol{a}$	Vector.
$\boldsymbol{A}$	Matrix.
$\boldsymbol{A}^{-1}$	Inverse of the matrix $\boldsymbol{A}$.
$\boldsymbol{A}^{\dagger}$	Pseudo-inverse of the matrix $\boldsymbol{A}$.
$\boldsymbol{A}^{T}$	Transpose of a matrix.
$\boldsymbol{A}^{*}$	Complex conjugate of a matrix.
$\boldsymbol{A}^{H}$	Transpose and complex conjugate of a matrix (Hermitian).
$\mathbf{0}$	Vector/matrix of zeros of the appropriate dimensions.
$\boldsymbol{I}$	Identity matrix of the appropriate dimensions.
$\mathrm{Tr}\left(\boldsymbol{A}\right)$	Trace of the matrix $\boldsymbol{A}$.
$\left\|\boldsymbol{A}\right\|$	Determinant of the matrix $\boldsymbol{A}$.
$\mathrm{rank}\left(\boldsymbol{A}\right)$	Rank of the matrix $\boldsymbol{A}$.
$\ker\left(\boldsymbol{A}\right)$	Kernel/null-space of the matrix $\boldsymbol{A}$.
$\mathrm{blkdiag}\left(\boldsymbol{A}_1, \dots, \boldsymbol{A}_N\right)$	Block diagonal matrix formed with the matrices $\{\boldsymbol{A}_1, \dots, \boldsymbol{A}_N\}$.
$\mathrm{diag}\left(a_1, \dots, a_N\right)$	Diagonal matrix whose main diagonal is $\{a_1, \dots, a_N\}$.
$\mathrm{eig}\left(\boldsymbol{A}\right)$	Diagonal matrix whose main diagonal are the eigenvalues of the matrix $\boldsymbol{A}$.
$\left\|\boldsymbol{a}\right\|_2$	Norm-2 (Euclidean norm) of the vector $\boldsymbol{a}$, see (3.49).
$\left\|\boldsymbol{A}\right\|_F$	Frobenius norm of the matrix $\boldsymbol{A}$, see (6.13).
$\boldsymbol{\nabla}_{\boldsymbol{x}} f(\boldsymbol{x})$	Gradient of a function $f(\boldsymbol{x})$, (3.53).
$\mathbb{E}\left\{\cdot\right\}$	Statistical expectation.
$\log\left(\cdot\right)$	Natural (base e) logarithm.
$\log_2\left(\cdot\right)$	Base 2 logarithm.

$[a]^+$	$\max(0, a)$.
$\mathbb{R}$	Field of real numbers.
$\mathbb{R}^+$	Set of positive real numbers.
$\mathbb{R}^+ \cup \{0\}$	Set of non-negative real numbers.
$\mathbb{C}$	Field of complex numbers.

Acronyms

3G 3^{rd} Generation of Mobile Communications.

3GPP 3^{rd} Generation Partnership Project.

4G 4^{th} Generation of Mobile Communications.

5G 5^{th} Generation of Mobile Communications.

AWGN Additive White Gaussian Noise.

BD Block Diagonalization.

BS Base Station.

CA Carrier Aggregation.

CDF Cumulative Distribution Function.

CoMP Coordinated Multi Point.

HetNet Heterogeneous Networks.

HSPA+ High-Speed Packet Access Plus.

iid independent identically distributed.

IMT-Advanced International Mobile Telecommunications-Advanced.

ITU-R International Telecommunication Union Radiocommunication Sector.

KKT Karush-Kuhn-Tucker.

LTE Long Term Evolution.

LTE-A Long Term Evolution Advanced.

MIMO Multiple Input Multiple Output.

MISO Multiple Input Single Output.

MMSE Minimum Mean Squared Error.

MRT Maximum Ratio Combining.

OCI Other Cluster Interference.

OFDM Orthogonal Frequency Division Multiplexing.

OFDMA Orthogonal Frequency Division Multiple Access.

PAPC Per Antenna Power Constraint.

PBPC Per Base Station Power Constraint.

pdf probability density function.

QoS Quality of Service.

RN Relay Node.

SINR Signal to Interference plus Noise Ratio.

SNR Signal to Interference Ratio.

SNR Signal to Noise Ratio.

SU Single User.

SVD Singular Value Decomposition.

TPC Total Power Constraint.

UFR Universal Frequency Reuse.

WCDMA Wide-band Code Division Multiple Access.

ZFBF Zero Forcing Beamforming.

Chapter 1

Introduction

Every new generation of cellular network technologies comes with a new set of requirements, dictated by the trends in the use of the mobile connectivity. A common requirement to every single generation is their striving for higher data rates and greater power efficiency. This motivates research and techology innovation in order to achieve the goals set for each generation.

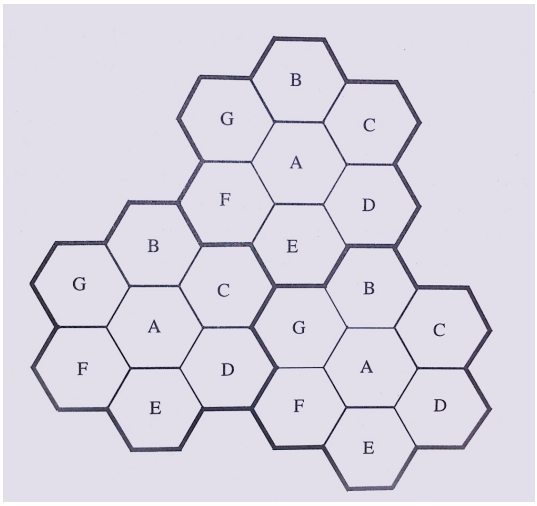
The research associated usually requires revisiting old paradigms used in previous generations, and updating them with novel ideas.

The *release 7* of the 3^{rd} Generation Partnership Project (3GPP) 3^{rd} Generation of Mobile Communications (3G) specifications [1], also known as High-Speed Packet Access Plus (HSPA+), included the use of Multiple Input Multiple Output (MIMO) as a means to increase the transmission rates.

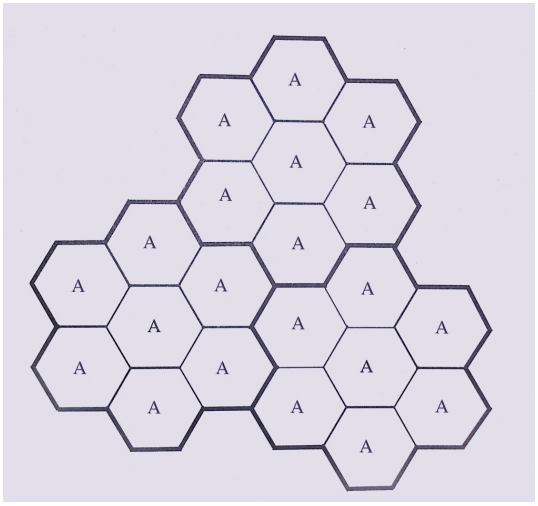
Release 8, more well known by its commercial name Long Term Evolution (LTE) [2], introduced a new physical layer, based on Orthogonal Frequency Division Multiplexing (OFDM) instead of Wide-band Code Division Multiple Access (WCDMA) as in 3G. Although the rates attainable with WCDMA may be comparable to those obtained with OFDM, the latter provides a much easier equalization mechanism that makes dealing with multipath channels a simpler task. Apart from that OFDM provides a higher flexibility in the resource allocation and user and enables the use of Orthogonal Frequency Division Multiple Access (OFDMA).

LTE did not meet the requirements issued by the International Telecommunication Union Radiocommunication Sector (ITU-R) International Mobile Telecommunications-Advanced (IMT-Advanced) radio interface [3] for what is known as 4^{th} Generation of Mobile Communications (4G) though.

The introduction of Long Term Evolution Advanced (LTE-A) in *release 10* of the LTE specification [4] met the requirements to be considered an IMT-Advanced system. The main novelties included in LTE-A are Carrier Aggregation (CA), enhanced use of MIMO techniques and support for Relay Nodes (RNs).



(a) Frequency reuse factor of $1/7$.



(b) Frequency reuse factor of 1. DO NOT FORGET TO NOT EMPHASIZE THE 7-CELL CLUSTERS BUT EACH CELL INDIVIDUALLY

Figure 1.1: Different frequency planning options

Release 11 [5] included in the specification the support for Coordinated Multi Point (CoMP) operation. CoMP was included in order to improve the network performance at cell edges, for it uses several transmitters to provide coordinated transmission in the downlink, and a number of receivers to provide coordinated reception in the uplink.

With LTE-A standardized and its deployment already ongoing, further releases of LTE-A still continue but standards bodies and industry are already looking ahead at the future 5^{th} Generation of Mobile Communications (5G), and so is doing the research world. Even though there is no definite idea about what 5G will be, it is clear what it will *not* be, an incremental advance on 4G. It needs to be a paradigm shift [6].

The new 5G systems will be characterized by being heterogeneous, what is known as Heterogeneous Networkss (HetNets), formed by multiple small cells, using different radio access technologies [7]. One of the main problems for HetNets is inter-cell interference, because of the possible presence of unplanned deployment of small cells, and the irregular shape of the cells. Hence the importance of interference coordination techniques.

Current MIMO systems used in cellular networks are not achieving the expected performance predicted by the initial theoretical works. The main reason for this is the interference that is present naturally in cellular systems when all cells share the same spectrum for the tranmissions. The effect of this interference is a reduction of the Signal to Interference plus Noise Ratio (SINR) experienced by the users, highly reducing the advantages that MIMO could potentially deliver.

The conventional approach for cellular networks was to perform a careful frequency planning in order to avoid the interference among neighboring cells. Clusters of N cells were

grouped together, and assigned N frequency bands to be used, and the pattern is repeated for different clusters, yielding what is called a *frequency reuse factor* of $1/N$, as exemplified in Figure 1.1a.

The problem that this poses is that the available spectrum must be split, which is an inherent inefficiency in the use of the resources.

A different option consists on a system where all the cells share a common spectrum, so that all of them can use the full amount of resources available. This is called Universal Frequency Reuse (UFR), and a graphical description can be seen in Figure 1.1b.

It is in this kind of networks that the need for coordination among cells arises, as every cell will interfere with the rest of the cells in the system reducing the SINR operating point of the users.

In the search for higher data rates and a more efficient use of the resources, UFR is a must to make the most out of the scarce resource that the radio frequency spectrum is. Therefore, "A new look at the interference" [8] is needed. The conventional concept of the interference as being an impairment needs to shift to a new point of view where the interference can be used to improve the overall performance of the network. A joint optimization of the resources among all the cells is required in order to globally improve the perfomance of the system [9].

The CoMP operation considered in [5] is just a part of a much broader field of multicell cooperation or coordinated communications where several cells are assumed to cooperate, in the sense that they take measures in order to alleviate to a certain degree the level of interference introduced into other parts of the network, or the use of that interference to their advantage.

Intuitively, the best strategy should be to allow all the Base Stations (BSs) in the network to cooperate, what is known as *global coordination*. Even though it may seem that global coordination may solve all the problems of frequency planning and resource allocation, it cannot be ignored that it comes at a non-negligible cost. The BSs in the network may need to interchange information in order to cooperatively transmit the information to all the users in the system. The amount of information that needs to be exchanged grows out of control with the size of the network, i.e., the number of BSs that form the system. The result of this is that the capacity required to transmit this information renders the alleged solution useless. Not only are the backhaul transmission capabilities required prohibitive, but also tight synchronization among the BSs becomes a challenge, and channel information gathering becomes a cumbersome task. Apart from this, theoretical works [10] have unveiled intrinsic limitations of cooperation, whose benefits do not unboundedly grow with the size of the coordination group.

For all these reasons, clustering appears as a means to cope with the limitations of global coordination. In clustering, the coordination is not performed among all the BSs in the network but, instead, small groups, or clusters, are formed and the cooperation takes place locally within the cluster. This greatly reduces the amount of control information that should be handled by the backhaul. Also, the reduced size of the group makes the system work at an

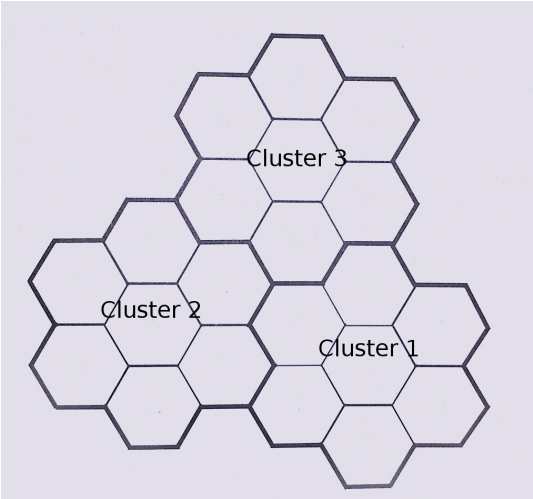


Figure 1.2: Clustered network scenario.

operating point where the natural limitations mentioned in [10] do not affect the performance of the network.

A schematic representation of a clustered network can be seen in Figure 1.2 where three clusters of seven cells are shown.

Grouping the cells in reduced size clusters has an important drawback: If the cooperation is done within a cluster and neighboring clusters are not coordinated in any way, there would be, again, unhandled interference, albeit not the same as in the uncoordinated scenario.

This thesis focuses on a clustered cellular network where Block Diagonalization (BD) is used for coordination within each cluster. The performance of the network, in terms of achievable rate and fairness considerations, is analyzed and its dependence on several parameters of the network is studied. Also, mechanisms to deal with the interference, resulting from clustering, are presented.

The organization of the document is as follows:

- In `\refc{ch:state_art}` a compilation of different alternatives for coordination, as well as for clustering, found in the literature are presented and described.
- `\refc{ch:system_model}` presents the system model used throughout the dissertation, and describes in detail BD and the power allocation strategies used in the rest of the work.
- `\refc{ch:achiev_rates}` analyzes the performance of a cellular network, in terms of the mean achievable rate as a function of the cluster size, when using BD for coordination within each cluster, and taking into account the interference due to external clusters. An analytical expression for the mean achievable rate is developed and the optimum cluster size is obtained.

- `\refc{ch:rate_statistics}` considers the fairness of the system, and studies the variability of the rate, as a complement to the mean obtained in `\refc{ch:achiev_rates}`. The behavior of the rates is shown to follow almost exactly a Gamma distribution.
- The pernicious effect of the Other Cluster Interference (OCI) in the rates is introduced in `\refc{ch:adaptive_schedule}`, and a mechanism to deal with it, based on a mixed tranmission strategy and on a scheduling algorithm, is presented.
- Finally, some conclusions are presented in `\refc{ch:conclusions}`, and future research topis are discussed.

Chapter 2

Analysis of the state of the art

Chapter 3

System Model

3.1 System Model

The system that will be considered throughout this work aims to represent the downlink of a canonical cellular network, comprising several identical cells, layed out over a regular hexagonal grid.

When studying a cellular network, the cells located at the edge of the network will not experience the same conditions as the cells in the center of the network. A typical way to deal with this situation is to consider a scenario that wraps around (Figure 3.2a) so that cells on one side of the scenario affect cells on the opposite side. Another option is to consider a scenario with more cells than necessary, and then analyze the behavior of the cells located within the center of the network (Figure 3.2b), so that the exterior cells account for the interference, equaling the conditions of all the cells in the network.

Each cell in the system under study will be served with a single BS that is equipped with t transmit antennas. Each of the users considered in the system has r receive antennas. Figure 3.1 shows a schematic representation of such a network, where R_{cell} is the cell radius, and d_{ij} is the distance from the j-th BS to the i-th user. Each user is supposed to be associated with a single BS and, without loss of generality, it is assumed that the i-th user is served by the i-th BS, and in the following the distance between a user and its serving BS will be denoted by $d_i \triangleq d_{ii}$.

A system with M BSs and N users can then be modelled as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{3.1}$$

where $\boldsymbol{y}$ represents the signal received at all the users and is defined as

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix} \in \mathbb{C}^{Nr \times 1} \tag{3.2}$$

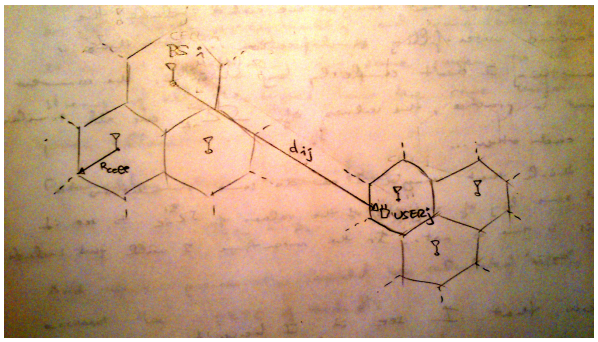
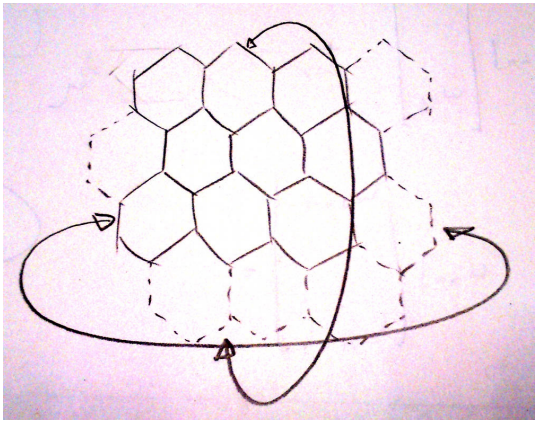
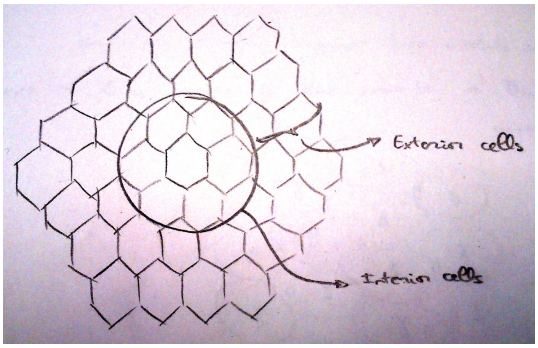


Figure 3.1: Schematic representation of the cellular network that will be used throughout this dissertation.



(a) Wrap around scenario. Cells on one side of the scenario influence the cells on the other side as if they were next to each other.



(b) Oversized scenario. The behavior of the network is analyzed in the central cells of the network, and the exterior cells compensate the network edge effects.

Figure 3.2: Different scenario configurations.

where $\boldsymbol{y}_i \in \mathbb{C}^{r\times 1}$ is the signal received at the i-th user.

$\boldsymbol{H}$ is the channel matrix representing the propagation from all the BSs to all the users, with the following structure

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{11} & \cdots & \boldsymbol{H}_{1M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N1} & \cdots & \boldsymbol{H}_{NM} \end{bmatrix} \in \mathbb{C}^{Nr\times Mt} \tag{3.3}$$

where $\boldsymbol{H}_{ij} \in \mathbb{C}^{r\times t}$ represents the channel matrix from the j-th BS to the i-th user. It will include the path loss due to propagation, small scale fading, shadowing, and any other characteristic of the radio channel that needs to be taken into consideration.

The vector $\boldsymbol{x}$ in (3.2) is the signal transmitted from all the BSs, and it is composed of

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_M \end{bmatrix} \in \mathbb{C}^{Mt \times 1} \tag{3.4}$$

where $\boldsymbol{x}_j \in \mathbb{C}^{t \times 1}$ is the signal transmitted by the j-th BS. Additionally, the power transmitted by the j-th BS can be calculated from the transmitted signal as

$$P_{j,\mathrm{tx}} = \mathrm{Tr}\left(\boldsymbol{x}_j \boldsymbol{x}_j^H\right) = \boldsymbol{x}_j^H \boldsymbol{x}_j \tag{3.5}$$

and each BS will have an independent power constraint

$$P_{j,\mathrm{tx}} \leq P_{j,\mathbf{max}}. \tag{3.6}$$

Finally $\boldsymbol{n}$ represents the Additive White Gaussian Noise (AWGN) at all the receivers

$$\boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_1 \\ \vdots \\ \boldsymbol{n}_N \end{bmatrix} \in \mathbb{C}^{Nr \times 1} \tag{3.7}$$

with $\boldsymbol{n}_i \in \mathbb{C}^{r \times 1}$ accounts for the Gaussian noise at the i-th receiver. Throughout this work $\boldsymbol{n}_i$ is considered to be formed by independent identically distributed (iid) entries, drawn from a zero mean, σ_i^2 variance Gaussian distribution, $\boldsymbol{n}_i \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_i^2 \boldsymbol{I}\right)$. The noise variance will be assumed the same for all the receivers.

In a general scenario, there may be cooperation among the BSs in the system so that the information intended for a particular user will be transmitted by several or all the BSs. Or, equivalently, each BS transmits a combination of the information of several users

$$\boldsymbol{x}_j = \boldsymbol{W}_{j1}^{(\mathrm{tx})} \boldsymbol{s}_1 + \cdots + \boldsymbol{W}_{jN}^{(\mathrm{tx})} \boldsymbol{s}_N \tag{3.8}$$

where $\boldsymbol{s}_i \in \mathbb{C}^{\ell \times 1}$ is the vector of information symbols to be transmitted to user i, with ℓ being the number of simultaneous symbols or streams to be transmitted to that user. $\boldsymbol{W}_{ji}^{(\mathrm{tx})} \in \mathbb{C}^{t \times \ell}$ is the precoding matrix used at the j-th transmitter for the data of the i-th user.

It will be assumed that the information symbols are independent and drawn from a Gaussian distribution such that $\boldsymbol{s}_i \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{s}_i}\right)$, where $\boldsymbol{R}_{\boldsymbol{s}_i} = \mathrm{diag}\left\{p_{i1}, \ldots, p_{i\ell}\right\} \in \mathbb{R}^{\ell \times \ell}$ contains the power allocated to each of the symbols in $\boldsymbol{s}_i$. The transmitted power can be expressed as

$$P_{j,\mathrm{tx}} = \sum_{i=1}^{N} \mathrm{Tr}\left(\boldsymbol{W}_{ji}^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_{ji}^{(\mathrm{tx}),H}\right). \tag{3.9}$$

The choice of the precoding matrices and of the receiving filter will determine the transmission strategy used. This Thesis will focus mainly on BD [11] which is described in Section 3.3.

On the receiver side, no cooperation among the users will be considered, so each user may perform, independently, additional processing of the received signal by applying a linear filter or equalizer

$$\hat{\boldsymbol{s}}_i = \boldsymbol{W}_i^{(\mathrm{rx})} \boldsymbol{y}_i \tag{3.10}$$

where $\boldsymbol{W}_i^{(\mathrm{rx})} \in \mathbb{C}^{r \times \ell}$ is the equalizer used at the i-th receiver.

Combining (3.1)–(3.8) it is possible to rewrite (3.1) as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{W}^{(\mathrm{tx})}\boldsymbol{s} + \boldsymbol{n} \tag{3.11}$$

where

$$\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_N \end{bmatrix} \in \mathbb{C}^{N\ell \times 1} \tag{3.12}$$

and the global precoding matrix is

$$\boldsymbol{W}^{(\mathrm{tx})} = \begin{bmatrix} \boldsymbol{W}_{11}^{(\mathrm{tx})} & \cdots & \boldsymbol{W}_{1N}^{(\mathrm{tx})} \\ \vdots & \ddots & \vdots \\ \boldsymbol{W}_{M1}^{(\mathrm{tx})} & \cdots & \boldsymbol{W}_{MN}^{(\mathrm{tx})} \end{bmatrix} \in \mathbb{C}^{Mt \times N\ell} \tag{3.13}$$

and it can be partitioned as

$$\boldsymbol{W}^{(\mathrm{tx})} = \left[\boldsymbol{W}_1^{(\mathrm{tx})}, \ldots, \boldsymbol{W}_N^{(\mathrm{tx})} \right] \tag{3.14}$$

with

$$\boldsymbol{W}_i^{(\mathrm{tx})} = \begin{bmatrix} \boldsymbol{W}_{11}^{(\mathrm{tx})} \\ \vdots \\ \boldsymbol{W}_{M1}^{(\mathrm{tx})} \end{bmatrix} \in \mathbb{C}^{Mt \times \ell}. \tag{3.15}$$

The channel matrix can then be partitioned as

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_N \end{bmatrix} \tag{3.16}$$

where $\boldsymbol{H}_i \in \mathbb{C}^{r \times Mt}$ is the channel matrix representing the propagation from all the BSs to the i-th user.

As it has already been mentioned, receiver cooperation is not going to be considered, so looking at a particular user, e.g., the i-th user, the signal that is received will be

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{W}^{(\mathrm{tx})} \boldsymbol{s} + \boldsymbol{n}_i \tag{3.17}$$

which can be rewritten as

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{W}_i^{(\mathrm{tx})} \boldsymbol{s}_i + \underbrace{\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{s}_j}_{\text{Interference}} + \boldsymbol{n}_i \tag{3.18}$$

so that it can be readily seen how other users' data appear as an interference term that degrades the received signal.

Defining the term of interference plus noise as

$$\boldsymbol{z}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{s}_j + \boldsymbol{n}_i \tag{3.19}$$

and using (3.18), the ergodic (mean) rate for the i-th user is given by [12], [13]

$$R_i = \mathbb{E}\left\{ \log_2 \left| \boldsymbol{I} + \boldsymbol{H}_i \boldsymbol{W}_i^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_i^{(\mathrm{tx}),H} \boldsymbol{H}_i^H \boldsymbol{R}_{\boldsymbol{z}_i}^{-1} \right| \right\} \tag{3.20}$$

where $\boldsymbol{R}_{\boldsymbol{z}_i} \in \mathbb{C}^{r \times r}$ is the covariance matrix of the noise plus the interference term in (3.18)

$$\boldsymbol{R}_{\boldsymbol{z}_i} = \boldsymbol{z}_i \boldsymbol{z}_i^H = \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} + \boldsymbol{n}_i \right) \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} + \boldsymbol{n}_i \right)^H . \tag{3.21}$$

In the rest of the work, there are a set of assumptions that will be made, mainly to guarantee the feasibility of some of the results obtained:

- The number of users will be the same as the number of BSs, i.e., $N = M$.
- The total number of antennas transmitting will be greater or equal than the total number of antennas at the receiver side, this is $Mt \geq Nr$.

- The number of streams transmitted to each user must be $\ell \leq r$, and in general it will be assumed the equality.

- There is no correlation, neither at the transmitters nor at the receivers, so that $\boldsymbol{H}$ is full rank or, equivalently, $\text{rank}\,(\boldsymbol{H}) = \min\,(Nr, Mt)$. As $Mt \geq Nr$, then $\text{rank}\,(\boldsymbol{H}) = Nr$.

- The power available at each of the BSs will be assumed the same, i.e., $P_{j,\max} = P_{\max}$ for all j.

3.2 Channel Model

As it has been said, each component of $\boldsymbol{H}_{ij}$ accounts for the propagation path loss, small scale fading, shadowing, and other characteristics of the channel.

In terms of propagation, the channel is typically decomposed as

$$\boldsymbol{H}_{ij} = \boldsymbol{R}_{\text{rx},ij}\boldsymbol{H}_{\text{iid},ij}\boldsymbol{R}_{\text{tx},ij} \tag{3.22}$$

where $\boldsymbol{R}_{\text{tx},ij} \in \mathbb{C}^{t\times t}$ is the spatial correlation at the transmitter, $\boldsymbol{H}_{\text{iid},ij} \in \mathbb{C}^{r\times t}$ is the uncorrelated fading channel, and $\boldsymbol{R}_{\text{rx},ij} \in \mathbb{C}^{r\times r}$ is the spatial correlation at the receiver. In the current work, unless stated otherwise, the channel is considered spatially uncorrelated, both at the transmitter and at the receiver, that is $\boldsymbol{R}_{\text{tx},ij} = \boldsymbol{I}$ and $\boldsymbol{R}_{\text{rx},ij} = \boldsymbol{I}$, as in [14].

The small scale characteristics considered are Rayleigh, so that the entries of $\boldsymbol{H}_{\text{iid},ij}$ are iid complex Gaussian random variables with zero mean and a variance given by the power path loss between the j-th BS and the i-th user.

The attenuation that the signal experiences due to the propagation varies according to an exponential power decay with exponent γ, so that the path loss[*] is calculated as

$$\text{pl}_{ij} = \text{pl}_0 \left(\frac{d_{ij}}{d_0}\right)^{-\gamma} \tag{3.24}$$

where pl_0 represents the attenuation at a reference distance d_0. For the analysis done in this work, but without any loss of generality, pl_0 and d_0 are assumed equal to 1.

[*] In natural units, although it is also commont to represent the path loss in decibels as

$$\text{PL}_{ij}\,(\text{dB}) \triangleq 10\log_{10}\left(\text{pl}_{ij}\right) \tag{3.23}$$

3.2.1 SNR

With the definition of the propagation model that is used in this work, it can also be explained the definition of SNR that is used for the theoretical analyses and the simulations.

Analogously to other works such as [15], the SNR, denoted as ρ, is defined with reference to the power received at the three-way corner of the cell, at a distance of R_{cell}, when the BS transmits at full power, so that the relationship between SNR and the noise power is given by

$$\rho = \frac{P_{\max} R_{\text{cell}}^{-\gamma}}{\sigma_n^2} \tag{3.25}$$

where it is assumed that all the BSs have the same maximum transmission power. Using (3.25) it is possible to calculate the noise power for a given SNR and vice versa.

3.3 Block Diagonalization

One possibility to cancel the inter-user interference is to diagonalize the channel matrix. Perfect diagonalization is only possible if $Mt \geq Nr$ [16], and it is achieved using the following precoding matrix

$$\boldsymbol{W}^{(\text{tx})} = \boldsymbol{H}^{\dagger}. \tag{3.26}$$

This solution is optimum only when every user has only one antenna. In the case under study, with multiantenna receivers, complete diagonalization of the channel matrix is suboptimal since each user is able to coordinate the processing of its received signal.

In [11] is stated that the optimum solution under the constraint that all inter-user interference be zero is obtained with $\boldsymbol{H}\boldsymbol{W}^{(\text{tx})}$ being block diagonal. In [11], BD is proposed as an algorithm to obtain a precoding matrix that is able to block diagonalize the channel matrix. This algorithm is described next.

In order to meet the condition of zero inter-user interference, it is necessary to cancel the interference term in (3.18), and this is equivalent to meet the following

$$\boldsymbol{H}_i \boldsymbol{W}_j^{(\text{tx})} = \boldsymbol{0} \quad \forall j \neq i. \tag{3.27}$$

Let $\widetilde{\boldsymbol{H}}_i$ be the channel matrix $\boldsymbol{H}$ with the rows corresponding to the matrix $\boldsymbol{H}_i$ removed, i.e.,

$$\widetilde{\boldsymbol{H}}_i = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_{i-1} \\ \boldsymbol{H}_{i+1} \\ \vdots \\ \boldsymbol{H}_N \end{bmatrix} \in \mathbb{C}^{(N-1)r \times Mt}. \tag{3.28}$$

Then the condition (3.27) can be obtained making $\boldsymbol{W}_i^{(\mathrm{tx})}$ lie in the null space or kernel of $\widetilde{\boldsymbol{H}}_i$. This is possible only if the dimension of the null space is greater than zero, i.e., $\operatorname{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) > 0$.

Now, with the dimensions of $\widetilde{\boldsymbol{H}}$, the rank of its null space is

$$\operatorname{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) = Mt - \operatorname{rank}\left(\widetilde{\boldsymbol{H}}_i\right). \tag{3.29}$$

But it is assumed that $\boldsymbol{H}$ is full rank, ergo $\operatorname{rank}\left(\widetilde{\boldsymbol{H}}_i\right) = (N-1)r = \widetilde{L}_i$, and then

$$\operatorname{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) = Mt - \widetilde{L}_i > 0 \tag{3.30}$$

so it is guaranteed that a precoding matrix $\boldsymbol{W}_i^{(\mathrm{tx})}$ that lies in the null space of $\widetilde{\boldsymbol{H}}_i$ exists.

The simplest way to obtain such $\boldsymbol{W}_i^{(\mathrm{tx})}$ involves using the Singular Value Decomposition (SVD) of the matrix $\widetilde{\boldsymbol{H}}_i$.

Let $\widetilde{\boldsymbol{H}}_i$ be decomposed as

$$\widetilde{\boldsymbol{H}}_i = \widetilde{\boldsymbol{U}}_i \widetilde{\boldsymbol{\Lambda}}_i \left[\widetilde{\boldsymbol{V}}_i^{(1)}, \widetilde{\boldsymbol{V}}_i^{(0)}\right]^H \tag{3.31}$$

where $\widetilde{\boldsymbol{V}}_i^{(0)} \in \mathbb{C}^{Mt \times (Mt - \widetilde{L}_i)}$ contains the last $Mt - \widetilde{L}_i$ right singular vectors of $\widetilde{\boldsymbol{H}}_i$, corresponding to the singular values equal to zero. $\widetilde{\boldsymbol{V}}_i^{(0)}$ forms an orthonormal basis of the null space of $\widetilde{\boldsymbol{H}}_i$, and thus its columns can be used to cancel the inter-user interference

$$\widetilde{\boldsymbol{H}}_i \widetilde{\boldsymbol{V}}_i^{(0)} = \boldsymbol{0}. \tag{3.32}$$

Using these matrices as precoding the result is

$$\widehat{\boldsymbol{H}} = \boldsymbol{H}\left[\widetilde{\boldsymbol{V}}_1^{(0)}, \ldots, \widetilde{\boldsymbol{V}}_N^{(0)}\right] = \begin{bmatrix} \boldsymbol{H}_1 \widetilde{\boldsymbol{V}}_1^{(0)} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{H}_N \widetilde{\boldsymbol{V}}_N^{(0)} \end{bmatrix} \tag{3.33}$$

that, as it can be seen, has a block diagonal structure, which gives the name to the algorithm proposed in [11].

The next problem that BD solves is the maximization of the sum-rate of the system, given the block diagonal structure in (3.33). The precoding matrix $\boldsymbol{W}_i^{(\mathrm{tx})}$ will be considered to be

$$\boldsymbol{W}_i^{(\mathrm{tx})} = \widetilde{\boldsymbol{V}}_i^{(0)} \boldsymbol{W}_i' \tag{3.34}$$

where $\boldsymbol{W}_i' \in \mathbb{C}^{(Mt-\widetilde{L}_i)\times \ell}$ will take care of the rate maximization.

Introducing (3.34) into (3.20) the ergodic capacity simplifies to

$$R_i^{\text{no interf}} = \mathbb{E}\left\{ \log_2 \left| \boldsymbol{I} + \frac{1}{\sigma_i^2} \widehat{\boldsymbol{H}}_i \boldsymbol{W}_i' \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_i'^{,H} \widehat{\boldsymbol{H}}_i^H \right| \right\} \tag{3.35}$$

where $\widehat{\boldsymbol{H}}_i = \boldsymbol{H}_i \widetilde{\boldsymbol{V}}_i^{(0)} \in \mathbb{C}^{r\times(Mt-\widetilde{L}_i)}$.

In order to maximize the rate, consider the SVD

$$\widehat{\boldsymbol{H}}_i = \widehat{\boldsymbol{U}}_i \begin{bmatrix} \widehat{\boldsymbol{\Lambda}}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \left[\widehat{\boldsymbol{V}}_i^{(1)}, \widehat{\boldsymbol{V}}_i^{(0)} \right]^H \tag{3.36}$$

where $\widehat{\boldsymbol{\Lambda}}_i = \mathrm{diag}\left\{ \hat{\lambda}_{i1}^{1/2}, \ldots, \hat{\lambda}_{ir}^{1/2} \right\} \in \mathbb{C}^{r\times r}$ contains the non-zero singular values of $\widehat{\boldsymbol{H}}_i$, which has rank $\left(\widehat{\boldsymbol{H}}_i\right) = r$. And $\widehat{\boldsymbol{V}}_i^{(1)} \in \mathbb{C}^{(Mt-\widetilde{L}_i)\times r}$ contains the first r right singular vectors of $\widehat{\boldsymbol{H}}_i$, and it will be used as $\boldsymbol{W}_i'$, yielding the following precoding matrix

$$\boldsymbol{W}_i^{(\mathrm{tx})} = \widetilde{\boldsymbol{V}}_i^{(0)} \widehat{\boldsymbol{V}}_i^{(1)} \tag{3.37}$$

The BD also provides the receiver filter to be used at each user which will be

$$\boldsymbol{W}_i^{(\mathrm{rx})} = \widehat{\boldsymbol{U}}_i^H \tag{3.38}$$

Using all of the above, the rate that the i-th user can obtain is given by the expression

$$R_i^{\mathrm{BD}} = \mathbb{E}\left\{ \sum_{k=1}^{\ell} \log_2 \left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2} \right) \right\} \tag{3.39}$$

where the only parameters left to be computed are the power allocated to each of the ℓ streams of each user, and different options to do it will be discussed in Section 3.4.

3.4 Power Allocation

In Section 3.3, the BD algorithm has been described to get the precoding matrix to be used at the transmitter and the equalization filter to be used at the receiver side. After BD has been used, the power should be allocated to each of the data streams of each user, this is, the p_{ij} in (3.39) should be calculated in order to achieve a given performance, and subject to particular constraints.

The need for a power allocation algorithm comes from the restriction on the maximum power available for transmission, which may be due to physical limitations, or regulatory issues.

There are several different power constraints:

- Per Antenna Power Constraint (PAPC): The maximum power is constrained for each antenna at the transmitter. This option is specially well suited for distributed antenna systems [17], [18].
- Per Base Station Power Constraint (PBPC): In this case the maximum power is limited per base station instead of per antenna. This option is more appropriate for scenarios where all the transmitting antennas are collocated and may share a power budget, so that the transmission power can be arbitrarily allocated to each of the transmitter antennas.

A Total Power Constraint (TPC) can also be considered, which assumes that the maximum power is shared among all the transmitters in the system. Although this system is more easily analyzed, it is very unrealistic so it will not be considered in this work, except in Subsection 3.4.3 where a TPC is used to obtain an intermediate result.

For the sake of simplicity, PBPC will be used for the different analyses. In any case, PAPC can be seen as a particularization of PBPC as the derivations shown in this section can be applied to a PAPC system considering instead of each BS to have t transmit antennas, t single antenna BSs.

The problem that needs to be solved is, in general, the maximization of some function of the rate of each of the users subject to a PBPC

$$\begin{aligned} \underset{\{\boldsymbol{R}_{\boldsymbol{s}_i}\}}{\text{maximize}} \quad & f\left(R_i^{\mathrm{BD}}\left(\boldsymbol{R}_{\boldsymbol{s}_1},\right),\ldots,R_N^{\mathrm{BD}}\left(\boldsymbol{R}_{\boldsymbol{s}_N}\right)\right) \\ \text{subject to} \quad & P_{j,\mathrm{tx}} \leq P_{j,\max}, \quad j=1,\ldots,M \end{aligned} \tag{3.40}$$

One common metric used for the maximization is the weighted sum-rate of the system, so that the function $f(\cdot)$ is equal to

$$f\left(R_i^{\mathrm{BD}},\ldots,R_N^{\mathrm{BD}}\right)=\sum_{i=1}^{N}\alpha_i R_i^{\mathrm{BD}} \tag{3.41}$$

where $\alpha_i \in [0,1]$ can be seen as different priorities for different users, and they are assumed to be

$$\sum_{i=1}^{N} \alpha_i = 1 \tag{3.42}$$

and in the particular case where all the $\alpha_i = 1/N$, then the function $f(\cdot)$ represents the sum-rate of the system.

Calling

$$\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} = \left[\boldsymbol{W}_{j1}^{(\mathrm{tx})}, \dots, \boldsymbol{W}_{jN}^{(\mathrm{tx})}\right] \in \mathbb{C}^{t \times N\ell} \tag{3.43}$$

the precoding matrix of the j-th BS, and

$$\boldsymbol{R}_{\boldsymbol{s}} = \mathrm{blkdiag}\left(\boldsymbol{R}_{\boldsymbol{s}_1}, \dots, \boldsymbol{R}_{\boldsymbol{s}_N}\right) \in \mathbb{C}^{N\ell \times N\ell} \tag{3.44}$$

the matrix containing the power assigned to all the streams of all the users, the power constraint in (3.40) can then be reformulated as

$$\mathrm{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right) \leq P_{j,\max} \tag{3.45}$$

Now the term inside the trace operator can be written explicitly as a function of p_{ik} in order to make it easier to analyze. First define

$$\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} = \left[\bar{\boldsymbol{w}}_{j,11}, \dots, \bar{\boldsymbol{w}}_{j,1\ell}, \dots, \bar{\boldsymbol{w}}_{j,N\ell}\right] \tag{3.46}$$

where $\bar{\boldsymbol{w}}_{j,ik} \in \mathbb{C}^{t\times 1}$ is the ik-th column of $\overline{\boldsymbol{W}}_j^{(\mathrm{tx})}$, i.e., the precoding that is used at the j-th BS for the k-th stream of the i-th user. And then:

$$\begin{aligned}
\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H} &= \left[\bar{\boldsymbol{w}}_{j,11}, \dots, \bar{\boldsymbol{w}}_{j,N\ell}\right] \begin{bmatrix} p_{11} & & 0 \\ & \ddots & \\ 0 & & p_{N\ell} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{w}}_{j,11}^H \\ \vdots \\ \bar{\boldsymbol{w}}_{j,N\ell}^H \end{bmatrix} = \\
&= p_{11} \bar{\boldsymbol{w}}_{j,11} \bar{\boldsymbol{w}}_{j,11}^H + \cdots + p_{N\ell} \bar{\boldsymbol{w}}_{j,N\ell} \bar{\boldsymbol{w}}_{j,N\ell}^H \\
&= \sum_{i=1}^{N} \sum_{k=1}^{\ell} p_{ik} \bar{\boldsymbol{w}}_{j,ik} \bar{\boldsymbol{w}}_{j,ik}^H
\end{aligned} \tag{3.47}$$

and the trace is

$$\mathrm{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})}\boldsymbol{R}_{\boldsymbol{s}}\overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right)=\sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 \tag{3.48}$$

where

$$\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2=\mathrm{Tr}\left(\bar{\boldsymbol{w}}_{j,ik}\bar{\boldsymbol{w}}_{j,ik}^H\right)=\bar{\boldsymbol{w}}_{j,ik}^H\bar{\boldsymbol{w}}_{j,ik}. \tag{3.49}$$

The sum-rate maximization problem can then be formulated in *standard form* [19] as

$$\begin{aligned}
\underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right) & & \\
\text{subject to} \quad & \sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2-P_{j,\max}\leq 0, & & j=1,\ldots,M \\
& -p_{ik}\leq 0, & & \begin{array}{l} i=1,\ldots,N \\ k=1,\ldots,\ell \end{array}
\end{aligned} \tag{3.50}$$

In the next sections, different alternatives for obtaining these powers are presented and described.

3.4.1 Optimal Power Allocation

In (3.50), the function $\log_2(\cdot)$ is convex on p_{ik} and the sum of convex functions is also convex, so the objective function in (3.50) is convex. The constraints are affine and therefore convex too. The optimization problem in (3.50) is a convex optimization problem that can be solved using a myriad of numerical technics [19]. Nonetheless, it would be interesting to analyze a bit further the problem in order to get some insight about it.

The problem in (3.50) satisfies *Slater's condition* [19] since the objective function is convex and all the inequality constraints are affine, hence *strong duality* holds. This means that the optimum value of the primal problem is equal to the optimum value of the *Lagrange dual problem*, so that this can be used to find out the solution to the primal, original, problem.

Under these conditions, and considering that the objective function is differentiable with respect to p_{ik}, Karush-Kuhn-Tucker (KKT) conditions [19] are necessary and sufficient for optimality of a solution, and they can be used to analyze the optimization problem in search for an optimal solution. The KKT conditions for (3.50) are

$$\begin{aligned}
\sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}^{*}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max} &\leq 0, \quad j=1,\dots,M \\
p_{ik}^{*} &\leq 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
\nu_j^{*} &\geq 0, \quad j=1,\dots,M \\
\mu_{ik}^{*} &\geq 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
\nu_j^{*}\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max}\right) &= 0, \quad j=1,\dots,M \\
-\mu_{ik}^{*}p_{ik}^{*} &= 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
\boldsymbol{\nabla}_{\boldsymbol{p}}\mathcal{L}\left(\boldsymbol{p}^{*},\boldsymbol{\nu}^{*},\boldsymbol{\mu}^{*}\right) &= \mathbf{0}
\end{aligned} \tag{3.51}$$

where the superscript * represents a feasible solution of the optimization problem, $\boldsymbol{\nabla}_{\boldsymbol{p}}$ is the gradient with respect to the powers $\boldsymbol{p} = [p_{11},\dots,p_{N\ell}]^T$, $\boldsymbol{\nu} = [\nu_1,\dots,\nu_M]^T$ and $\boldsymbol{\mu} = [\mu_{11},\dots,\mu_{N\ell}]^T$ are the Lagrange multipliers, and $\mathcal{L}$ represents the *Lagrangian* associated with the problem (3.50), and it is defined as

$$\begin{aligned}
\mathcal{L}\left(\boldsymbol{p},\boldsymbol{\nu},\boldsymbol{\mu}\right) = &-\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right) + \\
&\sum_{j=1}^{M}\nu_j\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max}\right) - \\
&\sum_{i=1}^{N}\sum_{k=1}^{\ell}\mu_{ik}p_{ik}
\end{aligned} \tag{3.52}$$

The gradient of a function $f(\boldsymbol{x})$ with respect to $\boldsymbol{x}\in\mathbb{C}^{n\times 1}$ is defined as

$$\boldsymbol{\nabla}_{\boldsymbol{x}}f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1}f(\boldsymbol{x}) \\ \vdots \\ \frac{\partial}{\partial x_n}f(\boldsymbol{x}) \end{bmatrix} \tag{3.53}$$

First the gradient of the objective function is calculated, by computing the partial derivatives

$$\frac{\partial}{\partial p_{ik}}\left\{-\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right)\right\} = \frac{-\hat{\lambda}_{ik}}{\log\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}\right)} \tag{3.54}$$

And the same for the inequality constraints

$$\begin{aligned} &\frac{\partial}{\partial p_{ik}} \left\{ \sum_{i=1}^{N} \sum_{k=1}^{\ell} p_{ik} \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 - P_{j,\max} \right\} = \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \\ &\frac{\partial}{\partial p_{ik}} \{p_{ik}\} = 1 \end{aligned} \tag{3.55}$$

So that the condition of the gradient of the Lagrangian vanishing, in (3.50) can be writen as

$$\frac{-\hat{\lambda}_{ik}}{\log(2)\left(\sigma_i^2 + \hat{\lambda}_{ik} p_{ik}^*\right)} + \sum_{j=1}^{M} \nu_j^* \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 - \mu_{ik}^* = 0, \qquad \begin{matrix} i = 1, \dots, N \\ k = 1, \dots, \ell \end{matrix} \tag{3.56}$$

It can be seen that μ_{ik} is a slack variable that takes into account the non-negativeness of the powers p_{ik}, and it can be ommited to get the equation

$$\frac{-\hat{\lambda}_{ik}}{\log(2)\left(\sigma_i^2 + \hat{\lambda}_{ik} p_{ik}^*\right)} + \sum_{j=1}^{M} \nu_j^* \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \geq 0, \qquad \begin{matrix} i = 1, \dots, N \\ k = 1, \dots, \ell \end{matrix} \tag{3.57}$$

Calling

$$L_{ik} = \sum_{j=1}^{M} \nu_j^* \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \tag{3.58}$$

(3.57) can be solved for p_{ik}^*

$$p_{ik}^* \leq \frac{1}{\log(2)\, L_{ik}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}}, \qquad \begin{matrix} i = 1, \dots, N \\ k = 1, \dots, \ell \end{matrix} \tag{3.59}$$

The result in (3.59) resembles the classical *water-filling* solution, except that now the water level is not fixed, and it depends on the precoders. The coupling existing among the power constraints of the different BSs makes it impossible to find a closed-form solution for the values of p_{ik}.

Nevertheless, this analysis motivates the development of suboptimal schemes that are described in the following sections.

3.4.2 Modified Water-Filling

[20] proposes a simplification to the original problem, in order to make it more tractable. The coupling of the power constraints in (3.50) makes it impossible to get a simple solution for

the optimal power allocation problem. [20] approaches the problem by first considering an equivalent virtual BS so that the problem is cast with a single power constraint.

In order to do so, instead of having a power constraint for each of the BSs consider a single power constraint given by the most restrictive BS in the original problem. Define

$$\Omega_{ik}^{\mathrm{BS}} = \max_{j=1,\dots,M} \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \tag{3.60}$$

as the weights of the single virtual BS corresponding to each of the users' streams. The optimization problem becomes then

$$\begin{aligned} \underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right) \\ \text{subject to} \quad & \sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\Omega_{ik}^{\mathrm{BS}} - P_{\mathrm{BS,max}} \leq 0 \\ & -p_{ik} \leq 0, \qquad\qquad i=1,\dots,N \\ & \qquad\qquad\qquad\quad k=1,\dots,\ell \end{aligned} \tag{3.61}$$

where $P_{\mathrm{BS,max}}$ represents the most restrictive power constraint among all of the BSs.

This problem meets the same conditions as the original problem so that a similar analysis can be used. First formulate the Lagrangian of the new problem as

$$\begin{aligned} \mathcal{L}\left(\boldsymbol{p},\nu,\boldsymbol{\mu}\right) = & -\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right) + \\ & \nu\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\Omega_{ik}^{\mathrm{BS}} - P_{\mathrm{BS,max}}\right) - \\ & \sum_{i=1}^{N}\sum_{k=1}^{\ell} \mu_{ik}p_{ik} \end{aligned} \tag{3.62}$$

And its gradient is given by

$$\boldsymbol{\nabla}_{\boldsymbol{p}}\mathcal{L}\left(\boldsymbol{p}^*,\nu^*,\boldsymbol{\mu}^*\right) = \frac{-\hat{\lambda}_{ik}}{\log\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}^*\right)} + \nu^*\Omega_{ik}^{\mathrm{BS}} - \mu_{ik}^* = 0, \qquad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \tag{3.63}$$

Using the KKT condition that the gradient of the Lagrangian should vanish, and considering μ_{ik}^* a slack variable, and solving for p_{ik}^*, the following inequality is obtained

$$p_{ik}^* \leq \frac{1}{\log\left(2\right)\nu^*\Omega_{ik}^{\mathrm{BS}}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}}, \qquad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \tag{3.64}$$

which, together with the constraint of the powers being non-negative, can be written as

$$p_{ik}^{*} = \left[\frac{1}{\log(2)\,\nu^{*}\Omega_{ik}^{\mathrm{BS}}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}} \right]^{+}, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \tag{3.65}$$

The solution of this simplified problem is given by the water-filling solution with a variable water level and, in this case, an uncoupled solution for each of the data streams for each user. This allows for the use of standard and efficient methods to find the power allocation [21].

Clearly, the definition of the new problem makes it more restrictive than the original, and its solution will be also a feasible solution for the original problem, albeit not the optimal. The results in [20] show how under some conditions, the solution achieved like this can be rather close to the optimum one.

Standard Water-Filling

One further simplification that is done in [20] is to consider that, in practice, the values of all the $\Omega_{ik}^{\mathrm{BS}}$ are very similar to each other, so that it is possible to consider them equal. This turns the problem into a standard water-filling problem, with constant water level, where the solution is given by

$$p_{ik}^{*} = \left[K_{\mathrm{WF}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}} \right]^{+}, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \tag{3.66}$$

where $K_{\mathrm{WF}} = \frac{1}{\log(2)\nu^{*}\Omega_{\mathrm{WF}}}$, is constant, as $\Omega_{ik}^{\mathrm{BS}} = \Omega_{\mathrm{WF}}\ \forall i = 1, \ldots, N;\ \forall k = 1, \ldots, \ell$ is the same for all the data streams of all the users.

3.4.3 Scaled Water-Filling

In [15] the same power allocation problem as in (3.50) is dealt with by considering a TPC, so that the optimization problem becomes

$$\begin{aligned} \underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log_2\left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2}\right) \\ \text{subject to} \quad & \mathrm{Tr}\left(\boldsymbol{W}^{(\mathrm{tx})}\boldsymbol{R}_{\boldsymbol{s}}\boldsymbol{W}^{(\mathrm{tx}),H}\right) - MP_{\max} \leq 0 \\ & -p_{ik} \leq 0, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \end{aligned} \tag{3.67}$$

where it has been assumed that $P_{j,\max} = P_{\max}\forall j$.

Under this TPC, the solution is readily derived by water-filling [21], but the resulting $\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}}$ may violate the individual power contraints of each BS.

In order to meet each PBPC, the matrix $\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}}$ must be scaled so that the final power allocation is given by

$$\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{SWF}} = \beta \boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}} \tag{3.68}$$

where the scaling factor $\beta \in (0, 1)$ is calculated as

$$\beta = \frac{P_{\max}}{\max\limits_{j=1,\ldots,M} \mathrm{Tr}\left(\boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}} \boldsymbol{W}_j^{(\mathrm{tx}),H}\right)} \tag{3.69}$$

The results in [15] show, as well, that this simplified approach can deliver near-optimum performance.

3.4.4 Uniform Power allocation

The simplest, both conceptually and computationally, alternative that can be considered to solve the power allocation in (3.50) consists on considering a uniform power allocation.

This approach assigns the same power to all the data streams of all the users. Formally this means

$$p_{ik} = p_s, \quad \begin{array}{l} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{array} \tag{3.70}$$

where the power p_s should be computed taking into account the PBPC for each BS.

Recall from (3.45) the power transmitted by the j-th BS, where now the matrix $\boldsymbol{R}_{\boldsymbol{s}}$ is given by

$$\boldsymbol{R}_{\boldsymbol{s}} = p_s \boldsymbol{I} \tag{3.71}$$

and (3.45) becomes

$$p_s \, \mathrm{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right) \leq P_{j,\max}, \quad j = 1, \ldots, M \tag{3.72}$$

The new power allocation problem can be formulated as

$$\begin{aligned} \underset{p}{\text{maximize}} \quad & p \\ \text{subject to} \quad & p \operatorname{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right) \leq P_{j,\max}, \qquad j = 1, \ldots, M \end{aligned} \tag{3.73}$$

which is a *linear programming* optimization problem, and it can be solved efficiently using classical methods, e.g., bisection method [22].

Chapter 4

Mean Achievable Rates in Clustered Coordinated Base Station Transmission with Block Diagonalization*

4.1 Introduction

In Chapter 2 it has been discussed how global coordination in a cellular network is not feasible for practical application. Research has shown that grouping cells in clusters may help to alleviate some of the problems of global coordination. The clustering solution is not unique, though, and multiple alternatives an numerous parameters have to be chosen in order to meet different objectives.

The objective of this chapter is to analyze the mean achievable rate in a cellular network where clusters have been formed, and where BD is used within each cluster to manage the interference.

With this analysis, further research can be made in order to be able to choose from one of the most important parameters in clustering, the cluster size.

Numerical simulations are also performed in order to validate the theoretical analysis developed.

4.2 Clustered Network

The network to be considered is organized as independent groups of M cells, each of which is a cluster where its cells coordinate in order to transmit to the N users that are located within

*The work shown in this chapter has been published in [23].

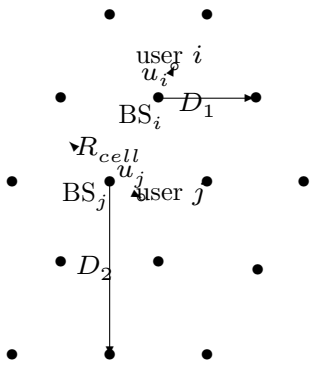


Figure 4.1: System layout with clusters of seven cells of radius R_{cell} (the radius of the circle circumscribing the cell) with an example of two users, with the distances $D_{\text{tier 1}}$ and $D_{\text{tier 2}}$, respectively, from their BS to the interfering BSs.

the cluster.

The clusters that form the network are non-overlapping, i.e., each cell in the system belongs to one and only one cluster. Overlapping, user-centric clusters have been shown to provide, in some situations, better performance than disjoint clusters [24], but this approach gives rise to a dramatic increase of the management complexity.

The clusters, then, are defined by the network planner, and they are kept fixed, grouping the BSs according to a distance criterion, so that the cells belonging to a cluster must form, borrowing the name from graph theory, a *connected component* of the graph containing all the cells in the system.

In the setup under study, all the BSs in the cluster are considered cluster members, that is, no scheduling or adaptive selection of active BSs is addressed in this work. In any case, they could be considered as a special case of the optimization problem to obtain the power allocation scheme.

Figure 4.1 shows such a clustered network, where the disjoint clusters can be observed, together with other parameters that will be used in considering the interference for the analysis.

4.3 Interference Model

In Figure 4.1 an example of a network is shown, with three complete clusters of $M = 7$ cells each, and with some cells belonging to clusters not completely shown.

The user i, at a distance d_i from its serving BS, *cf.* Section 3.1, is affected by the interference originated in the neighboring clusters. In this case, the closest interfering cells are located at a distance $D_{\text{tier 1}} = \sqrt{3}R_{\text{cell}}$ from the i-th BS.

Similarly, for the user j, at a distance d_j from its serving BS, *cf.* Section 3.1, the nearest interfering cells are located at a distance $D_{\text{tier 2}} = 3R_{\text{cell}}$ from the j-th BS.

Due to the cellular geometry, for a cluster size of up to 18, only these two possibilities exist: the closest interfering cell is at a distance of either $D_{\text{tier 1}}$ or $D_{\text{tier 2}}$ from the serving BS of each user.

The hexagonal cell can be approximated by a circular one with radius R_{cell}, see Figure 4.1, and then, assuming a uniform distribution of the users over each cell, the probability density function (pdf) of the distance of a user to its serving (closest) BS is given by

$$f_{d_i}(d_i) = \frac{2d_i}{R_{\text{cell}}^2} \tag{4.1}$$

The interference power received at the user i is equal to

$$I_i(d_i) = \sum_{m=1}^{M_{\text{interf}}} P_{\max} \hat{d}_{im}^{-\gamma} \tag{4.2}$$

where M_{interf} is the number of interfering BSs, i.e., the total number of cells in the system minus the M cells that form the cluster, and $\hat{d}_{im}$ is the distance from the m-th BS outside the cluster to the i-th user in the cluster. All the interfering BSs are assumed to be transmitting at full power.

In order to simplify the computation of the interference power, an equivalent model is introduced. In this model, the interference comes from $M_{\text{eq},i}$ cells, all of which are located at the same distance $(D_i - d_i)$ from the i-th user, the one being interfered.

The distance D_i takes the value of the distance from the serving BS to the closest interfering BS

$$D_i \in \{D_{\text{tier 1}}, D_{\text{tier 2}}\} \tag{4.3}$$

The equivalent number of interfering BSs, $M_{\text{eq},i}$, is such that the total interference power is the same as in the original layout.

With all this, (4.2) can be written as

$$I_i(d_i) = P_{\max} M_{\text{eq},i} (D_i - d_i)^{-\gamma} \tag{4.4}$$

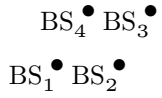


Figure 4.2: Cluster with $M = 4$ cells and neighbor interfering cells.

This approach is similar to that followed in [25], where a fluid model network is used. This model assumes that there is a continuum of BSs interfering, but this will not be considered in the current work for the sake of simplicity.

The real and equivalent model produce the same total interference, provided that $M_{\mathrm{eq},i}$ is adequately selected.

In order to determine $M_{\mathrm{eq},i}$ the only interference that is accounted for is the one coming from the first tier of neighboring cells. This implies that different cluster configurations may have different number of interfering BSs for each of the cells in the cluster, and this number for the i-th cell is denoted as $M_{\mathrm{int},i} \leq M_{\mathrm{interf}}$.

This is made clear in Figure 4.2 where a cluster with $M = 4$ cells is surrounded by $M_{\mathrm{interf}} = 12$ cells. In this cluster, cells 1 and 3 experience an interference coming from $M_{\mathrm{int},1} = M_{\mathrm{int},3} = 4$ neighboring cells, while cells 2 and 4 receive the interference from $M_{\mathrm{int},2} = M_{\mathrm{int},4} = 3$ cells.

In order to approximate the value of equivalent interfering BSs for a general network setup, fist consider a simple scenario Figure 4.3 with a cluster of $M = 1$ cells, where a single user, $i = 1$, in the cell is affected by an interference power I_1 coming from all the $M_{\mathrm{int},1} = 6$ belonging to the first tier.

An assumption that can be made is that half of the $M_{\mathrm{int},1}$, i.e., three, BSs are located at a distance $(D_1 - d_1)$ and the other half are located at a distance $(D_1 + d_1)$. Accordingly, the interference power can be expressed as

$$I_1(d_1) = P_{\max}\left[\frac{3}{(D_1 + d_1)^{\gamma}} + \frac{3}{(D_1 + d_1)^{\gamma}}\right] \tag{4.5}$$

which is equal to (4.4) if the equivalent number of interfering BSs is defined as

$$M_{\mathrm{eq},1} = 3\left[1 + \left(\frac{D_1 - d_1}{D_1 + d_1}\right)^{\gamma}\right] \tag{4.6}$$

In general, if a user i is considered to have $M_{\mathrm{int},i}$ interfering BSs in the first tier at a distance D_i, defined as in (4.3), then the equivalent number of interfering BSs is given by

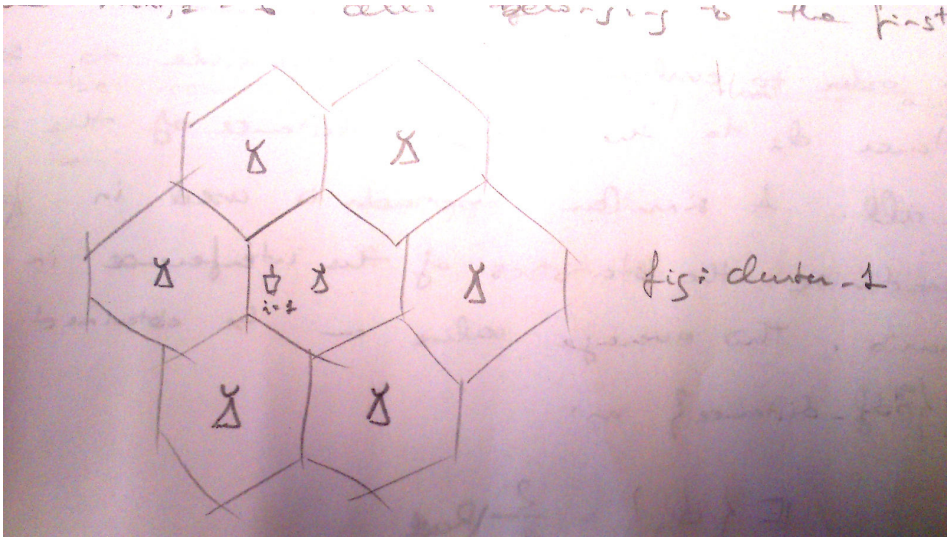


Figure 4.3: Simple scenario with a single cluster of $M = 1$ cell, and the six interfering cells surrounding it.

$$M_{\mathrm{eq},i} \approx \frac{M_{\mathrm{int},i}}{2}\left[1+\left(\frac{D_i-d_i}{D_i+d_i}\right)^{\gamma}\right] \tag{4.7}$$

In order to evaluate (4.7), it is possible to set the distance d_i to the average distance of the user within the cell. A similar approach is used in [26] to characterize the statistics of the interference in a multicell scenario. This average distance can be obtained from the uniform spatial distribution in (4.1) as

$$\mathbb{E}\{d_i\} = \frac{2}{3}R_{\mathrm{cell}} \tag{4.8}$$

and then

$$M_{\mathrm{eq},i} \approx \frac{M_{\mathrm{int},i}}{2}\left[1+\left(\frac{D_i-\frac{2}{3}R_{\mathrm{cell}}}{D_i+\frac{2}{3}R_{\mathrm{cell}}}\right)^{\gamma}\right] \tag{4.9}$$

Section 4.5 deals with the comparison between simulations and the analytical results, showing that the approximation in (4.9) is accurate.

4.4 Analysis of the Rate

In the case of having global coordination or, equivalently, having only one cluster including all the BSs, the interference among the users is completely eliminated through the use of BD, *cf.*

Section 3.3. On the other hand, in a multicluster environment, it is necessary to consider the effect of the interference coming from the cells outside the cluster. Hence, the mean achievable rate in (3.39) becomes (dropping the expectation notation)

$$R_i^{\mathrm{BD}} = \sum_{k=1}^{\ell} \log_2 \left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2 + I_i} \right) \tag{4.10}$$

where the parameter I_i represents the average power of the total interference contributions received in each data stream of user i from the interfering BSs.

It can be seen in (4.10) that the rate depends on the distance d_i of each user's equipment from the center of its cell. Using the pdf in (4.1) the average of the rate over all possible locations, making explicit the dependence on the distance d_i, can be expressed as

$$\begin{aligned} \bar{R}_i^{\mathrm{BD}} &= \int_0^{R_{\mathrm{cell}}} R_i^{\mathrm{BD}}(u) \frac{2u}{R_{\mathrm{cell}}^2} \mathrm{d}u \\ &= \int_0^{R_{\mathrm{cell}}} \sum_{k=1}^{\ell} \log_2 \left(1 + \frac{\hat{\lambda}_{ik}(u) p_{ik}}{\sigma_i^2 + I_i(u)} \right) \frac{2u}{R_{\mathrm{cell}}^2} \mathrm{d}u \end{aligned} \tag{4.11}$$

In (4.11) there are three parameters that determine the overall average rate, namely

- The interference $I_i(d_i)$ coming from outside the cluster.
- The effect of the channel fading and of the path loss, represented by the term $\hat{\lambda}_{ik}(d_i)$.
- The power p_{ik} assigned to the k-th data stream of user i.

In the following, the characterization of each of these parameters will be approached separatedly.

4.4.1 Interference

As described in Section 4.3 the contribution of interference, $I_i(d_i)$ on each data stream of user i, coming from the cells outside the cluster, can be considered as generated by an equivalent number of BSs located all of them at a distance of $D_i - d_i$ from the user. Recall that, for clusters of size up to 18, D_i can take one of two values as in (4.3).

The interfering distance D_i is then normalized to the cell radius R_{cell} by setting

$$\bar{d}_i = \frac{D_i}{R_{\mathrm{cell}}} \tag{4.12}$$

In expression (4.10) it is assumed that the interference can be treated as Gaussian noise, so that the power is calculated as the variance of that noise. Throughout the simulations that were performed, it has been found that, on average, at least 25 out-of-cluster cells contributed with significant interference.†

This number of 25 is dependent on the simulation paremeters, but it gives an idea of the order of magnitud of interferers present, and it justifies the treatment of the interference as Gaussian noise, by virtue of the central limit theorem [27].

4.4.2 Fading Effect

The terms $\hat{\lambda}_{ij}$ are the squared diagonal values of the matrix $\widehat{\boldsymbol{\Lambda}}_i$, *cf.* Section 3.3. This matrix is obtained in (3.36) by the combination of the channel matrix $\boldsymbol{H}_i$ and a unitary matrix, $\widetilde{\boldsymbol{V}}_i^{(0)}$.

The channel matrix $\boldsymbol{H}_i$ is composed of the submatrices $\boldsymbol{H}_{ij}$, where the fading elements have a power path loss of $d_{ij}^{-\gamma}$, and the elements of $\boldsymbol{H}_{ij}$ are independent from the elements of $\boldsymbol{H}_{ik}$ for all j different from k.

It is possible to define an alternative set of coefficients κ_{ij} that do not include the path loss effect

$$\kappa_{ij} \triangleq \frac{\hat{\lambda}_{ij}}{d_i^{-\gamma}} \tag{4.13}$$

These coefficientes are the elements of the main diagonal of the matrix

$$d_i^{\gamma} \widehat{\boldsymbol{\Lambda}}_i \widehat{\boldsymbol{\Lambda}}_i^H \tag{4.14}$$

And it can be seen that these diagonal elements are the singular values of the matrix

$$d_i^{\gamma} \boldsymbol{H}_i \widetilde{\boldsymbol{V}}_i^{(0)} \widetilde{\boldsymbol{V}}_i^{(0),H} \boldsymbol{H}_i^H \tag{4.15}$$

where $\boldsymbol{H}_i$ has Gaussian entries, and $\widetilde{\boldsymbol{V}}_i^{(0)}$ is a unitary matrix.

In the case of having $Mt = Nr$, the coefficientes κ_{ij} are the eigenvalues of a Wishart matrix while, in the general case of $Mt \geq Nr$ the matrix in (4.15) can be approximated by a Wishart matrix. Through simulations, it has been verified that the mean of the eigenvalues of both matrices, the original in (4.15) and the approximate Wishart, is the same and the difference between the two Cumulative Distribution Function (CDF) is less than 10 %.

The joint pdf of the eigenvalues κ_{ij} of a Wishart matrix can be obtained when the columns of the corresponding Gaussian matrix have an identity covariance matrix

†Being significant defined as being greater than the power received at the cell edge.

$$\boldsymbol{\Sigma}_i = \boldsymbol{I} \tag{4.16}$$

and it is given by [28]

$$f_{\kappa_{i1},\dots,\kappa_{i\ell}}(\kappa_{i1},\dots,\kappa_{i\ell}) = e^{-\sum_{k=1}^{\ell}\kappa_{ik}} \prod_{k=1}^{\ell} \frac{1}{[(\ell-k)!]^2} \prod_{m=k+1}^{\ell} (\kappa_{im}-\kappa_{i\ell})^2 \tag{4.17}$$

However, in the evaluation of the rate the complete pdf is not needed, and only the sum

$$\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log(\kappa_{ik}) \tag{4.18}$$

is required, which represents the expectation of the logarithm of the determinant, for which results are available, also for the general case when the covariance matrix is different from the identity, $\boldsymbol{\Sigma} \neq \boldsymbol{I}$, and it is given by [28]

$$\mathbb{E}\left\{\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log(\kappa_{ik})\right\} = \sum_{m=1}^{N\ell} \psi(N\ell-m) + \log(|\boldsymbol{\Sigma}|) \tag{4.19}$$

where $\psi(\cdot)$ is the Euler's digamma function [29], and where the matrix $\boldsymbol{\Sigma}$ is

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}_N \end{bmatrix} \tag{4.20}$$

with $\boldsymbol{\Sigma}_i$ the covariance matrix of the columns of the matrix $d_i^{\gamma/2}\boldsymbol{H}_i\widetilde{\boldsymbol{V}}_i^{(0)}$

$$\boldsymbol{\Sigma}_i = \boldsymbol{I}\left[1 + \sum_{\substack{j=1 \\ j\neq i}}^{M} \left(\frac{d_i}{d_{ij}}\right)^{\gamma}\right] \tag{4.21}$$

It is possible to define a parameter G_i

$$G_i \triangleq 1 + \sum_{\substack{j=1 \\ j\neq i}}^{M} \left(\frac{d_i}{d_{ij}}\right)^{\gamma} \tag{4.22}$$

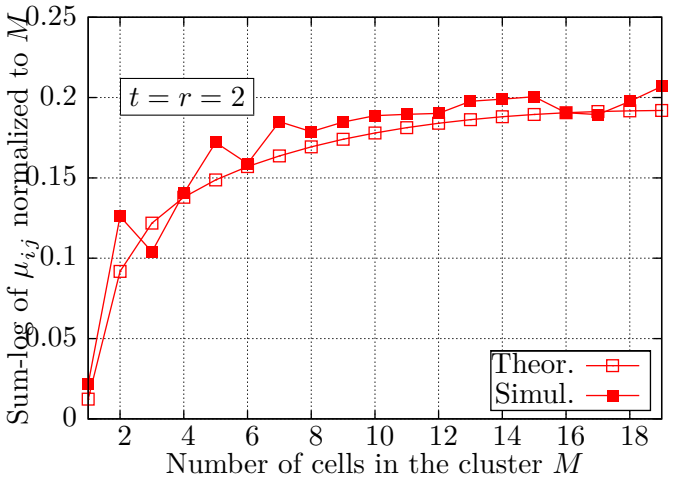


Figure 4.4: Sum-log of the terms κ_{ij}. Comparison between simulations and the values obtained using (4.19) for $t = r = 2$.

that can be considerd as a cluster gain. Its value can be approximated considering the average distance of a user to its BS (4.8), and the distance from the rest of the BSs in the cluster to be either $D_{\text{tier 1}}$ or $D_{\text{tier 2}}$. If additionally these two distances are normalized by the average distance in (4.8)

$$\begin{aligned} \overline{D}_{\text{tier 1}} &\triangleq \frac{3D_{\text{tier 1}}}{2R_{\text{cell}}} \\ \overline{D}_{\text{tier 2}} &\triangleq \frac{3D_{\text{tier 2}}}{2R_{\text{cell}}} \end{aligned} \tag{4.23}$$

then the gain factor G_i can be approximated for different cluster sizes as

$$G_i = \begin{cases} 1 + \frac{M-1}{2}\left[\left(\frac{1}{\overline{D}_{\text{tier 1}}-1}\right)^{\gamma}\left(\frac{1}{\overline{D}_{\text{tier 1}}+1}\right)^{\gamma}\right] & M \leq 7 \\ 3\left[\left(\frac{1}{\overline{D}_{\text{tier 1}}-1}\right)^{\gamma}\left(\frac{1}{\overline{D}_{\text{tier 1}}+1}\right)^{\gamma}\right]\frac{M-7}{2}\left[\left(\frac{1}{\overline{D}_{\text{tier 2}}-1}\right)^{\gamma}\left(\frac{1}{\overline{D}_{\text{tier 2}}+1}\right)^{\gamma}\right] & 7 < M \leq 18 \end{cases} \tag{4.24}$$

In Figure 4.4 a fixed distance for the users equal to $d_i = {}^2/_3 R_{\text{cell}}$, (4.8), is considered and the values of the sum of the natural logarithm of the values κ_{ij} are calculated through simulations and using the expression (4.19), for the case of $t = r = 2$.

It can be seen how the sum of the log values of κ_{ij} presents a diminishing increase as the cluster size M increases. This can be explained by a reduced contribution of the BSs, that are farther than in smaller clusters, which becomes negligible due to the path loss.

Notice that, in the evaluation of the mean achievable rate, a factor ${}^1/_N$ is applied in order to evaluate the average rate per user, taking into consideration that in a cluster with more cells, there would be more users as well.

Thus, a decrease occurs in the mean achievable rate per user for large values of M, as it will be shown in Section 4.5.

4.4.3 Power allocation

Under the BD strategy, the transmission within each cluster is equivalent to a set of parallel non-interfering channels.

Therefore, the transmission power must be allocated in order to optimize some quality of service parameters, such as the sum-rate or a weighted sum of the rates, for the users of each cluster.

This objective is subject to a maximum transmission power available at each BS, (3.45) and (3.48)

$$\sum_{i=1}^{N} \sum_{k=1}^{\ell} p_{ik} \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \leq P_{\max} \tag{4.25}$$

for each of the M BSs in the cluster.

The rate maximization problem is described in more detail in Section 3.4, and the solutions range from the simplest uniform power allocation, *cf.* Subsection 3.4.4, to an optimal allocation, *cf.* Subsection 3.4.1. In any case, the problem of power allocation is not the focus of this work because it can be solved separately and the actual powers could be inserted in the analytical expressions developed.

Hence, in the following a theoretical framework is derived for the uniform power allocation scheme, for the sake of simplicity, and an example with a different power allocation will be presented with the results.

With a uniform power allocation a common average power p_s is used for every stream of every user, as seen in (3.70). This value p_s varies according to the number of BSs in the cluster, decreasing for a larger size of the cluster, since a fraction of the overall available power is spent in the coordination, to null the interference.

Substituting all the p_{ik} in (4.25) by the common value p_s it is easy to see that the condition in (4.25) is limited by the BS for which the following factor is maximum

$$\chi_j \triangleq \sum_{i=1}^{N} \sum_{k=1}^{\ell} \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \tag{4.26}$$

Assuming that the coefficients of the precoding matrix, i.e., the elements of the vector $\bar{\boldsymbol{w}}_{j,ik}$, are Gaussian then χ_j is a Chi-squared random variables with $N' \triangleq N\ell t$ degrees of freedom, and then the power p_s is related to the reciprocal value of the maximum of M random variables

$$p_s = \frac{P_{\max}}{\mathbb{E}\{\chi\}} \tag{4.27}$$

with

$$\chi = \max\{\chi_1, \dots, \chi_M\} \tag{4.28}$$

Then the probability distribution function of χ is given by

$$F_\chi(x) = P(N', x)^M \tag{4.29}$$

where $P(\cdot,\cdot)$ is the regularized Gamma function.

The mean value can be derived from the probability distribution function in (4.29) as

$$\mathbb{E}\{\chi\} = \int_0^\infty \left(1 - F_\chi(x)\right) \mathrm{d}x \tag{4.30}$$

and it can be bounded using

$$(1 - e^{-\alpha x})^a \leq P(a, x) \leq (1 - e^{-\beta x})^a \tag{4.31}$$

with

$$\begin{aligned} \alpha &= \begin{cases} 1 & 0 < a < 1 \\ \Gamma(a+1)^{-\frac{1}{a}} & a > 1 \end{cases} \\ \beta &= \begin{cases} \Gamma(a+1)^{-\frac{1}{a}} & 0 < a < 1 \\ 1 & a > 1 \end{cases} \end{aligned} \tag{4.32}$$

where $\Gamma(\cdot)$ is the Gamma function.

The average value of χ is then bounded by

$$\frac{1}{\beta}\left[\psi(Mt+1) + \gamma_0\right] \leq \mathbb{E}\{\chi\} \leq \frac{1}{\alpha}\left[\psi(Mt+1) + \gamma_0\right] \tag{4.33}$$

with $\psi(\cdot)$ again the Euler's digamma function, and γ_0 the Euler-Mascheroni constant.

It is possible to rewrite (4.33) in terms of p_s as

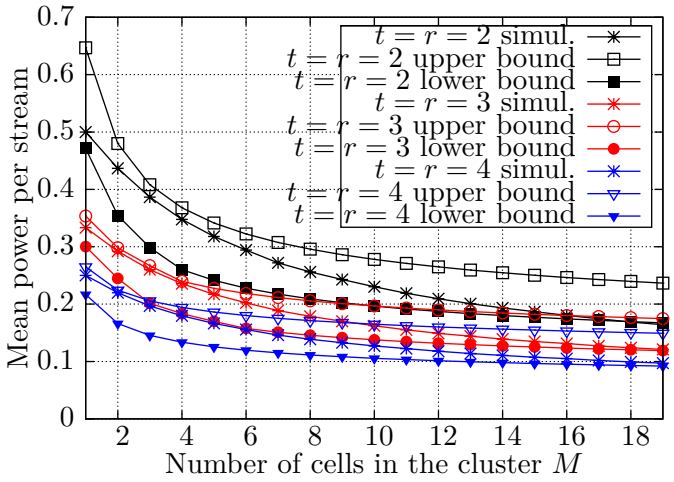


Figure 4.5: Normalized average power per stream $p_s/P_{\max}$ for different antenna configurations: comparison between simulations and the bounds in (4.34).

$$P_{\max}\frac{\Gamma\left(N'+1\right)^{-\frac{1}{N'}}}{\psi\left(Mt+1\right)+\gamma_0} \leq p_s \leq P_{\max}\frac{1}{\psi\left(Mt+1\right)+\gamma_0} \tag{4.34}$$

In the evaluation of the rate, the lower bound in (4.34) will be considered, providing a lower bound to the average rate of each user.

The bounds for the power per stream p_s in (4.34) are compared in Figure 4.5 with the results obtained through simulations.

First thing that can be seen in Figure 4.5 is how the power p_s decreases with the size of the cluster, and this affects the mean achievable rate as it will be discussed in Section 4.5.

Secondly, Figure 4.5 shows a very good agreement between the analytical and simulation results for different antenna configurations. In particular, the upper bound is tight for small clusters while the lower bound becomes more accurate for bigger clusters.

4.4.4 Evaluation of the mean achievable rate

The performance of the coordination scheme will be measured by the mean achievable rate per user in the cluster

$$\bar{R}^{\mathrm{BD}} = \frac{1}{N}\sum_{i=1}^{N}\bar{R}_i^{\mathrm{BD}} \tag{4.35}$$

It is possible to derive a lower bound for each user's average rate in (4.11) by considering the inequality

$$\log(1+x) \geq \log(x) \tag{4.36}$$

so that the average rate for the i-th user (4.11) becomes

$$\begin{aligned} \bar{R}_i^{\mathrm{BD}} &\geq \frac{1}{\log(2)} \sum_{k=1}^{\ell} \int_0^{R_{\mathrm{cell}}} \log\left(\frac{p_{ik}\kappa_{ik} u^{-\gamma}}{\sigma_i^2 + P_{\max} M_{\mathrm{eq},i} \left(D_i - u\right)^{-\gamma}} \right) \frac{2u}{R_{\mathrm{cell}}^2} \mathrm{d}u \\ &= \frac{1}{\log(2)} \sum_{k=1}^{\ell} \left\{ \log(\kappa_{ik}) + \log\left(\frac{p_{ik}}{P_{\max}} \right) + Z_i \right) \end{aligned} \tag{4.37}$$

where the interference model has been introduced, and Z_i is defined as

$$Z_i \triangleq \int_0^{R_{\mathrm{cell}}} \log\left(\frac{u^{-\gamma}}{\frac{\sigma_i^2}{P_{\max}} + M_{\mathrm{eq},i} \left(D_i - u\right)^{-\gamma}} \right) \frac{2u}{R_{\mathrm{cell}}^2} \mathrm{d}u \tag{4.38}$$

The value of $Z_i$is derived in Appendix A, and it is

$$\begin{aligned} Z_i = {} & \frac{\gamma}{2} + \log(\rho) + \bar{d}_i^2 \log\left(\frac{\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}} \right) \\ & - 2\bar{d}_i^2 \gamma \, {}_2\mathrm{F}_1\left(1, \frac{1}{\gamma}; \frac{\gamma+1}{\gamma}; \frac{-\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}} \right) \\ & + \frac{\bar{d}_i^2}{2} \gamma \, {}_2\mathrm{F}_1\left(1, \frac{2}{\gamma}; \frac{\gamma+2}{\gamma}; \frac{-\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}} \right) \\ & - \left(\bar{d}_i^2 - 1\right) \log\left(\frac{\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho + \left(\bar{d}_i - 1\right)^{\gamma}} \right) \\ & + 2\bar{d}_i \left(\bar{d}_i - 1\right) \gamma \, {}_2\mathrm{F}_1\left(1, \frac{1}{\gamma}; \frac{\gamma+1}{\gamma}; \frac{-\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho} \right) \\ & - \frac{\left(\bar{d}_i - 1\right)^2}{2} \gamma \, {}_2\mathrm{F}_1\left(1, \frac{2}{\gamma}; \frac{\gamma+2}{\gamma}; \frac{-\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho} \right) \end{aligned} \tag{4.39}$$

where ${}_2\mathrm{F}_1(\cdot)$ is the hypergeometric function, and ρ is the SNR as defined in (3.25).

Combining (4.19), (4.34), (4.35), and (4.37) the analytical expression for the mean achievable rate per user is

Figure 4.6: Scenario used for the simulations, with 7 hexagonal tiers.

$$\begin{aligned}\bar{R}^{\mathrm{BD}} \geq \frac{1}{N\log(2)} \Bigg[& \log\left(|\boldsymbol{\Sigma}|\right) + \sum_{m=1}^{N\ell} \psi\left(N\ell - m\right) \\ & + N\ell \log\left(\frac{\Gamma\left(N'+1\right)^{-\frac{1}{N'}}}{\psi\left(Mt+1\right)+\gamma_0}\right) + \ell \sum_{i=1}^{N} Z_i \Bigg]\end{aligned} \tag{4.40}$$

4.5 Numerical Results

In this section, the results derived from the analytical expression in (4.40) are compared with the results obtained through simulations.

The scenario for the simulations is a network composed of 169 cells, laid out as 7 concentric tiers of hexagonal cells Figure 4.6.

The cell radius, unless stated otherwise, is assumed to be $R_{\mathrm{cell}} = 1.4\,\mathrm{km}$.

All the results are averaged over 5,000 random trials. In each of these trials the position of the users was randomly set according to a uniform distribution inside each cell, Section 4.3. Also, for each of the trials, a random channel was generated according to the model described in Section 3.2, with a path loss exponent of $\gamma = 3.8$.

The parameter evaluated in the simulations is the achievable rate defined in (4.10), in which the different variables required (transmission power, interference power, $\widehat{\lambda}_{ik}$, etc) were obtained by simulations.

As it has already been mentioned, the clusters considered are static, i.e., they are fixed and do not change for all the simulations. Despite this, not all the clusters must have the same shape for the same cluster size M. In fact, in the simulations, the clusters were generated following a heuristic approach that tries to group cells in a compact way, with a regular shape, by minimizing the sum of the inter-cell distances, thus to avoid long clusters. Note,

Figure 4.7: Irregular shaping of the clusters, due to the heuristic clustering algorithm used.

however, that some values of M do not allow for regular clusters, i.e., hexagonal, to be formed. Figure 4.7 shows an example of this situation, where not all the clusters have the same shape.

4.5.1 Analytical and simulation results comparison

In order to validate the analytical expression (4.40), first it is compared with the mean achievable rate simulated using BD and, although the analytical derivations did not consider it and for the sake of completeness, the Minimum Mean Squared Error (MMSE) precoder described in [30].

The antenna configuration used for this first comparison is $r = t = 2$, and also different values of the SNR, as defined in (3.25), are used so to observe the behaviour in different SNR regimes.

Figure 4.8 shows the mean achievable rate as a function of the cluster size M.

As expected, the MMSE approach outperforms the BD strategy at low SNR. On the other hand, for moderate values of the SNR BD is able to provide comparable, and even more favorable, results, thus showing that the interference dominates over the noise for regimes other than the low SNR regime.

A very good agreement between the theoretical result in (4.40) and the simulations is clear in Figure 4.8, where also some variations can be seen in the simulation results. This is mainly due to the variability of the cluster shape, as seen in Figure 4.7, for different cluster sizes, and not to a low number of simulations that were averaged. That irregular shape of the clusters, despite being more or less controlled in the heuristic cluster selection algorithm, affects the simulation results in the form of the variability shown in the figures.

[10] points out at a fundamental limit of cooperation, and it is shown how the gains from cooperation cannot be unbounded, and that increasing the number of coordinated elements

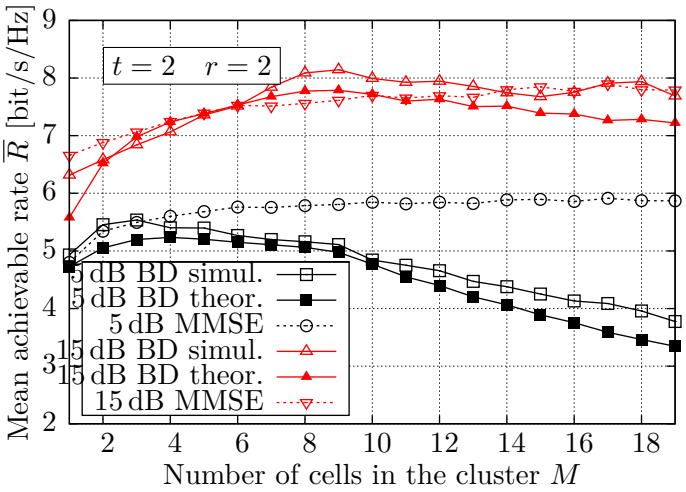


Figure 4.8: Mean achievable rate per user as a function of the number of cells in the cluster for $r = 2$, $t = 2$, variable values of SNR, $\gamma = 3.8$.

may saturate the performance achieved. This very same behaviour can be observed in Figure 4.8, both for BD and MMSE, where the mean achievable rate do not grow unboundedly with the cluster size, and an optimum value of the size M can be found.

Figure 4.9 also shows a similar behaviour. In this case, the mean achievable rate is represented as a function of the SNR, and there is an SNR at which the rate stops growing. This threshold SNR depends on the propagation path loss exponent γ because it directly determines the influence of the interference. In particular, the saturation SNR for BD is higher than for MMSE. For the former, it is always above 20 dB, for the scenarios considered, even for very small path loss exponents. This means that the saturation occurs for relatively large values of the SNR which is of practical importance because it would be possible to deliver good performance, using BD, within a practical range of SNR values.

Under the restriction that the same $P_{\max}$ is transmitted for all values of γ, the saturation occurs at different levels for each γ, although the general conclusions do not change.

In order to complete the validation of the theoretical results with the simulations, a fixed value of SNR= 25 dB and different antenna configurations were considered in Figure 4.10, still using uniform power allocation. The discrepancies between the theoretical results and the simulations are not only due to the approximations, but also to the fact that in the simulation scenario not all the clusters have the same shape, despite having the same number of cells.

4.5.2 Effect of the power allocation

The cause of the decrease of the rate with respect to the cluster size, as seen in Figure 4.8, is two-fold:

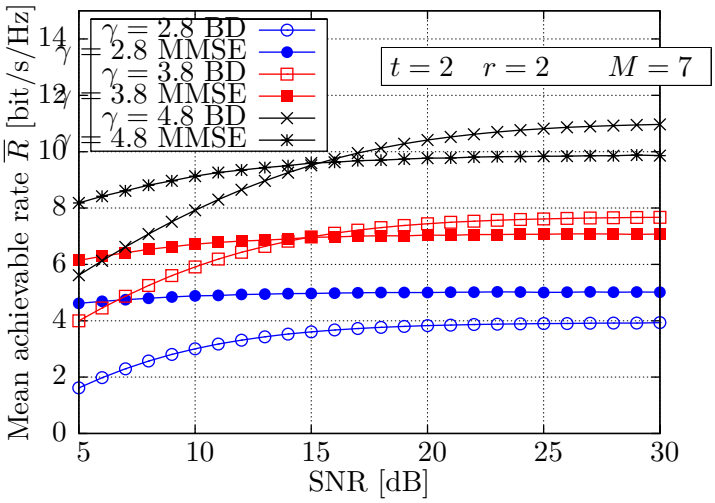


Figure 4.9: Mean achievable rate as a function of the SNR for different values of the path loss coefficient γ.

- First, the value of the attenuation experienced by each data stream, $\hat{\lambda}_{ik}$ decreases when the size of the cluster increases, as shown in Figure 4.4.
- Second, the power assigned to each data stream, the terms p_{ik} that for a uniform power allocation are all equal to p_s, also decreases as the the cluster grows, Figure 4.5, due to a coordination "loss".

In this section, a different power allocation scheme, other than the uniform, was used in order to verify if the behavior observed in Figure 4.8, where the rate decreases with the cluster size, is due to the power allocation.

The optimal power allocation used was obtained by means of numerical optimization, as described in [20], using CVX [31], [32], and those powers were plugged into (4.40) instead of the uniform power allocation.

In Figure 4.11 the rates obtained with the uniform and the optimum power allocation are compared and represented versus the size of the cluster.

Although, as expected, the optimal outperforms the uniform power allocation, both curves show a similar trend, meaning that a reduced growth of the rate (and in some situationes a reduction of the rate itself) is not due to the power allocation scheme, but it is the manifestation of a fundamental limitation of the coordination scheme, along the lines of the results in [10].

4.5.3 Optimum cluster size

Figure 4.8 through Figure 4.11 show a common trend which is the rate having a reduced growth, or even a reduction, with the cluster size.

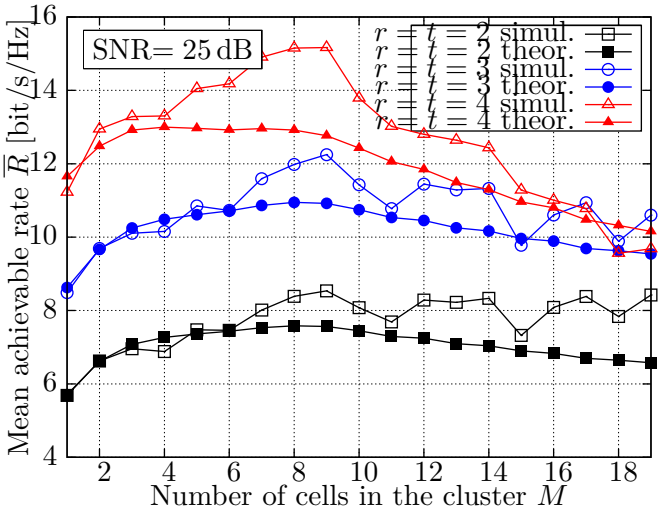


Figure 4.10: Mean achievable rate per user as a function of the cluster size for SNR= 25 dB, $\gamma = 3.8$ and different antenna configurations.

Given this, it is possible and interesing to find the cluster size M that can maximize the mean achievable rate.

In the case of the rate actually decreasing with M, the optimum value can be readily obtained as the value of M for which the maximum rate is obtained.

In some of the simulations results, the rate does not decrease within the range of cluster sizes that were simulated, so it is not possible, with the simulation conditions used, to find a maximum for the mean achievable rate. Something that can be observed, nevertheless, is that its growth with M is diminished so that the optimum value of M can be calculated by considering the relative change of the rate. The optimum is assumed to be found when the marginal increase of the rate is bellow a given percentage. In the case under study, the threshold was set to a 10 %.

Figure 4.12 represents the optimum value of M as a function of the SNR, for different antenna configurations, and for the power allocation schemes considered until now, uniform and optimum.

It can be seen that, for a wide range of SNR, the optimum value is limited to around 7–10 cells.

Only for high SNR is it more convenient to increase the cluster size, since the reduction of the interference can compensate the decrease of the cluster gain due to the decrease of the factors $\hat{\lambda}_{ik}$. This only happens for the case of considering the optimal power allocation, because in the case of the uniform power scheme there is the additional decrease of the power allocated to each stream, p_s.

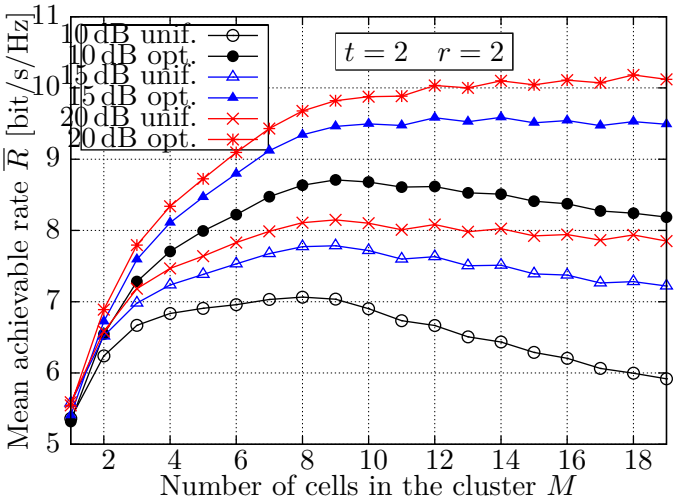


Figure 4.11: Comparison between different power allocation schemes, namely uniform and optimal. Mean achievable rate per user as a function of the cluster size for $r = t = 2$, $\gamma = 3.8$ and different values of SNR.

4.5.4 Effect of signalling overhead

It is common in the literature to not take into account the effect of the signalling overhead, and the same has been done in all the previous results of this work.

However, if a certain percentage of the available resources are dedicated to channel estimation and control signalling, the effective SNR and the payload that can be delivered are reduced with respect to the global achievable rate.

In [10] the overhead, incurred by channel estimation requirements, is accounted for by a percentage α which should grow, at least, linearly with the cluster size, up to a maximum value, to prevent it from being greater than 100 %. In that work, the SNR and the rate were effectively reduced by a factor of α, getting $\mathtt{SNR}_{\text{eff}} = (1 - \alpha)\,\mathtt{SNR}$ and $R_{\text{eff}} = (1 - \alpha)\,R$.

In order to show the effect of the overhead on the achievable rate, in Figure 4.13 a very conservative approach is adopted, in which the value of the reduction factor α scales linearly with the cluster size M, up to a maximum of 10 % for a cluster size of 19.

Figure 4.13 shows the comparison of considering and not considering the overhead. It can be seen that even with a small amount of overhead, increasing values of M lead to a worse performance.

Moreover, it should be stressed that the actual definition of signalling overhead and its management is usually delegated to the operator implementation, and this is seldom defined in the standards. Thus, its quantitative effect can change considerably depending on how the overhead is defined.

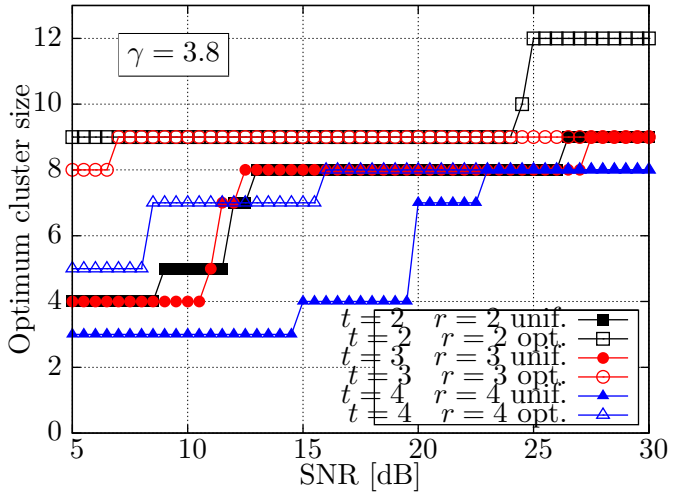


Figure 4.12: Optimum cluster size as a function of the SNR, $\gamma = 3.8$ and different antenna configurations.

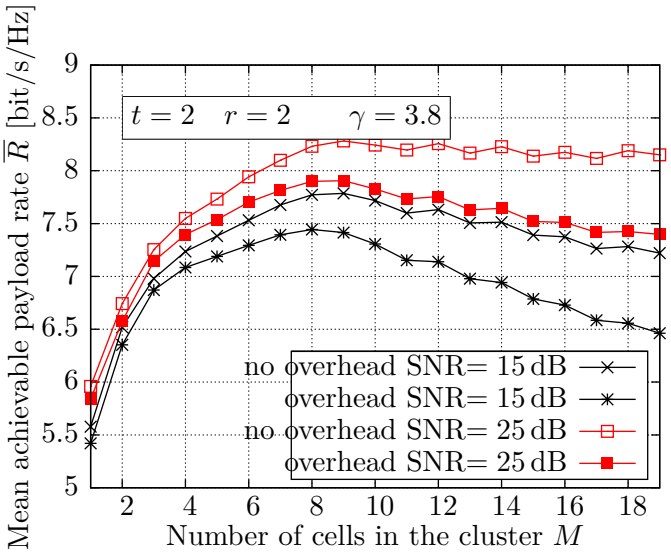


Figure 4.13: Mean achievable rate and payload rate as a function of the cluster size M with different SNRs.

Chapter 5

Achievable Rate and Fairness in Coordinated Base Station Transmission*

5.1 Introduction

Chapter 4 deals with the evaluation of the mean achievable rate, and the analysis of the influence of the cluster size and power assignment on the maximization of the rate. This chapter in turn focuses on the analysis of the Quality of Service (QoS) of the system, in terms of the fairness in the distribution of the achievable rate among the users.

Although the mean achievable rate, as calculated in Chapter 4 can provide useful information, also the CDF plays an important role when designing fairness, and other QoS related, management strategies in coordinated downlink networks.

In the following, it is shown that the statistics of the achievable rate are almost perfectly represented by a Gamma distribution. Note that the Gamma distribution arises in several contexts when considering non-negative random variables whose value is determined by several joint distributions. For example, in [34] it is used to model the interference in an interference-limited cellular network in order to simplify the calculation of the success probability, i.e., the complementary event of an outage, and the ergodic rate. In [35] a Gamma distribution is used to model composite fading channels. And in [36] a mixture Gamma distribution is proposed to describe the SNR of wireless channels, mainly composite shadowing/fading channels, but also many other small-scale fading channels.

This chapter provides a characterization of the CDF of the achievable rate in a coordinated base station transmission with BD, where the BSs are grouped in clusters and the interference is due to adjacent clusters.

*The work shown in this chapter has been published in [33].

Using the system model from Chapter 3, including the power assignment strategies described there, and the cluster model proposed in Section 4.2, the dependence of the CDF on the cluster size, and the power assignment strategy is studied as well in the current chapter.

5.2 Rate Statistics

5.2.1 Cumulative Distribution Function

Recal the expression of the achievable rate when BD is used to coordinate the BSs in the cluster

$$R_i^{\mathrm{BD}}(d_i) = \sum_{k=1}^{\ell} \log_2\left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2 + I_i(d_i)}\right) \tag{5.1}$$

where the dependence on the i-th user's distance to its BS, d_i, is made explicit, the interference power is defined by (4.2), the noise power σ_i^2 is obtained as a function of the SNR using (3.25), and the powers p_{ik} would be computed using the power allocation strategies described in Section 3.4.

The derivation of the complete statistical characterization of the achievable rate per user is an almost intractable task, due to the combination of many effects such as the power assignment, the interference coming from outside the cluster, the channel characteristics, etc. Even the evaluation of the mean requires to resort to several approximations, although the final result has been shown to be quite accurate, *cf.* Chapter 4.

If the pdf of the achievable rate per user is considered, a suitable analytical model is provided by the Gamma distribution's pdf

$$f_R(x) = \frac{1}{\Gamma(\theta)\,\theta^k} x^{k-1} e^{-\frac{x}{\theta}} \tag{5.2}$$

with mean and variance given by $k\theta$ and $k\theta^2$, respectively, and related to the system parameters as discussed in Subsection 5.2.3 and Subsection 5.2.4.

The choosing of the Gamma distribution is not arbitrary, and it is motivated by the fact that the achievable rate per user in (5.1) is the result of adding several non-negative terms. The *central limit theorem for causal functions* [27] states that the convolution of an unbounded number of causal functions can be approximated using a Gamma distribution. A causal function is a function defined only for $\mathbb{R}^+ \cup \{0\}$, as $\log_2(1+x)$ with $x \geq 0$ is. In the case under study the sum of terms is not unbounded, but [27] shows also how the approximation is accurate also for a sum of a finite number of terms.

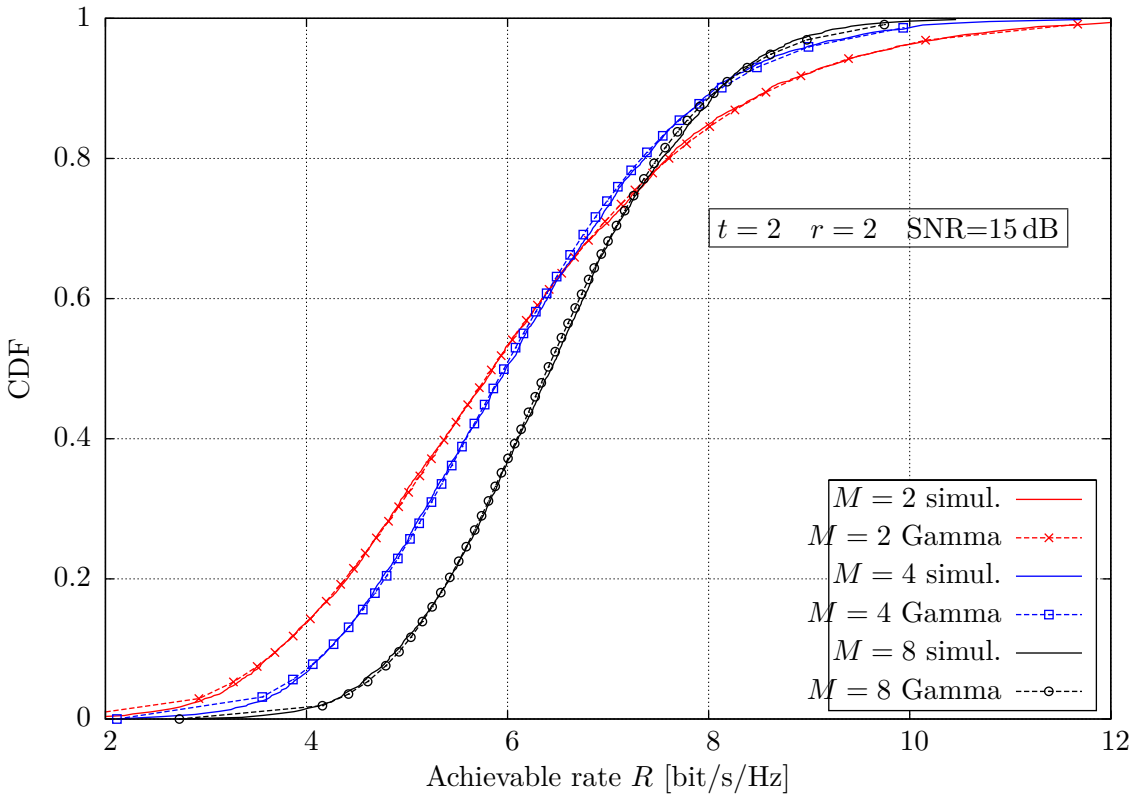


Figure 5.1: CDF of the achievable rate per user with $t = r = 2$ and SNR=15 dB. Comparison with the Gamma distribution with the same mean and variance.

The Gamma distribution has been introduced also in [25] to describe the Signal to Interference Ratio (SNR) in a simpler environment, without noise and without any kind of coordination.

In fact, when comparing the CDF obtained by simulation and a Gamma CDF, with the same mean and variance, it is interesting to note a very accurate fitting.

Figure 5.1 shows the perfect match between the experimental CDF and a Gamma CDF for a uniform power allocation, and for different system parameters.

The effect of increasing the size of the cluster on the CDF of the achievable rate is shown in Figure 5.2

Figure 5.2 also shows how the rate increases with the cluster size up to a certain point, and then it starts to decrease. This is the same behavior observed in the mean achievable rate in Chapter 4.

Another aspect that can be seen is how the variance of the rate decreases as the cluster size increases.

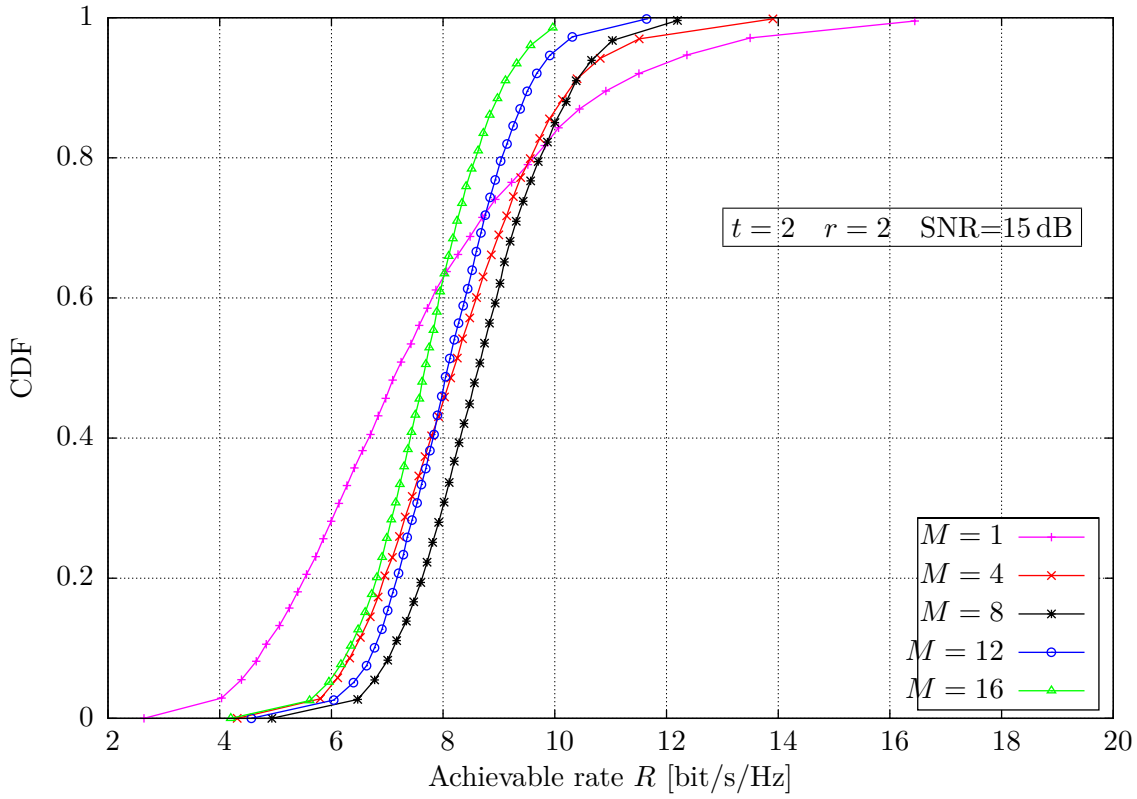


Figure 5.2: CDF of the achievable rate per user with $t = r = 2$ and SNR=15 dB for different values of the cluster size.

5.2.2 Effect of the power allocation

Three different power allocation schemes have been considered:

- Uniform power allocation, as described in Subsection 3.4.4.
- The modified water-filling from Subsection 3.4.2.
- Power allocation result of solving the convex optimization problem (3.50) using numerical solvers. In particular the numerical solver used is CVX [31], [32]. For the rest of the chapter, this solution will be noted as CVX solution.

In order to show the effect of using each of these power allocation schemes, Figure 5.3 presents the CDF of the achievable rate per user for the three of the schemes proposed, and using a constant antenna setup of $t = r = 3$ antennas, an SNR=15 dB, and two different cluster sizes, namely $M \in \{5, 8\}$.

It should be noted that both the water-filling and the numerical solutions are calculated for the problem formulated in (3.50), where no out-of-cluster interference is present, so that these

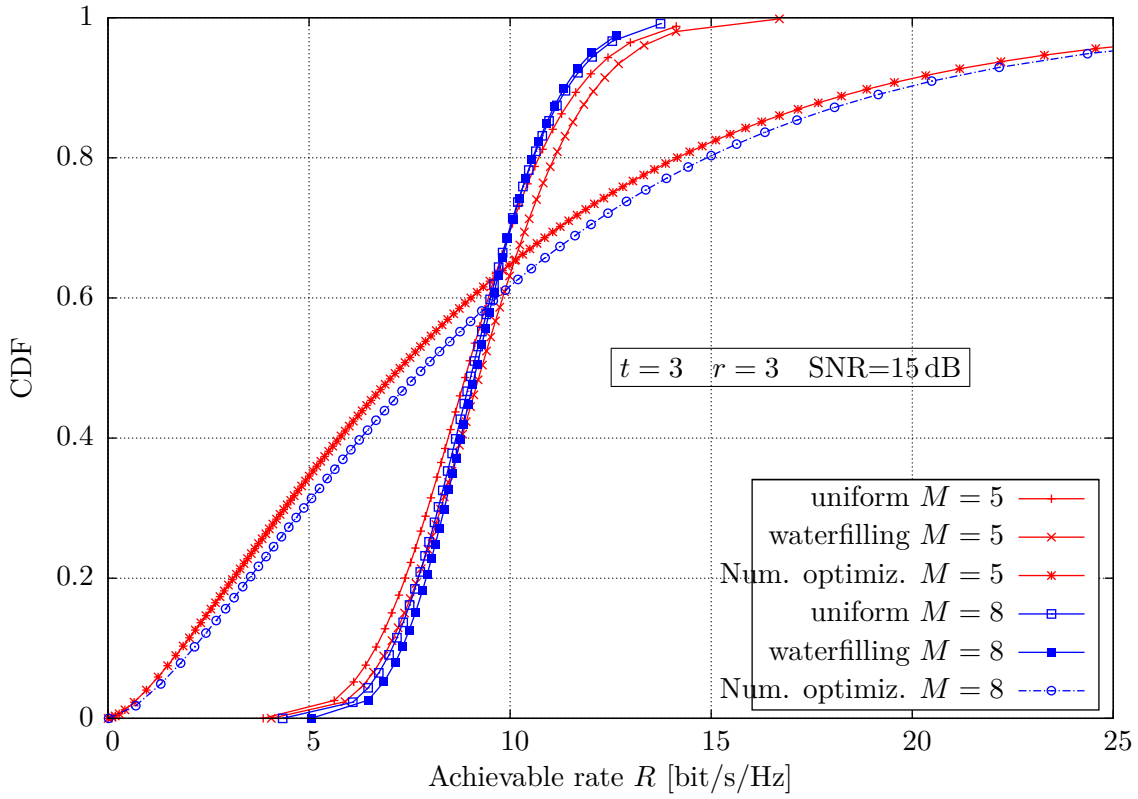


Figure 5.3: CDF of the achievable rate per user with $t = r = 3$ and SNR=15 dB. Comparison between different power allocation schemes.

solutions are not adapted to a multicluster environment. This is so because such adaptation would require some sort of coordination among different clusters, and that is out of the scope of this work.

5.2.3 Mean value

The derivation of the mean value has been done in Chapter 4 resorting to some approximations that allow to evaluate the mean achievable rate defined as in (4.11)

$$\bar{R}_i^{\mathrm{BD}} = \int_0^{R_{\mathrm{cell}}} R_i^{\mathrm{BD}}(u) \frac{2u}{R_{\mathrm{cell}}^2} \mathrm{d}u \tag{5.3}$$

using a particular model for the interference coming from outside the cluster, *cf.* Section 4.3.

The value of the mean that is thus obtained is rather accurate for a wide range of the scenario parameters, as it is shown in detail in Chapter 4.

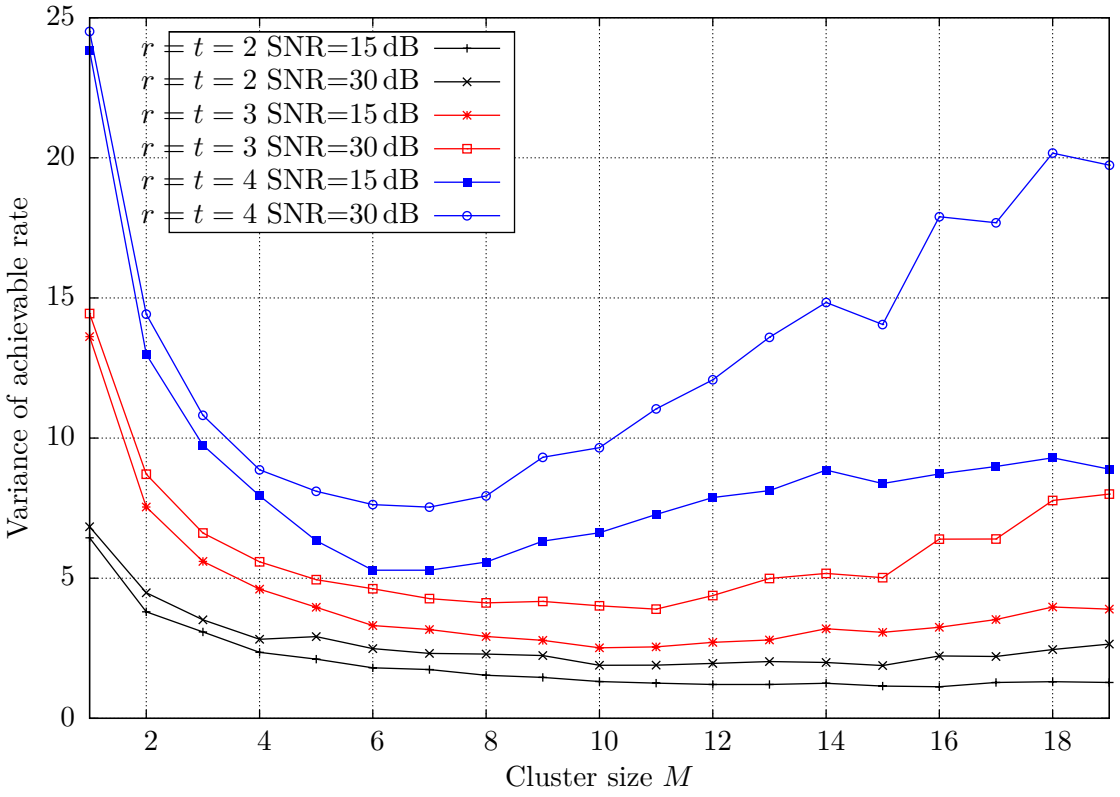


Figure 5.4: Variance of the achievable rate per user with different values of SNR and of t and r as a function of the cluster size M.

5.2.4 Variance

A closed form expression for the variance of the achievable rate cannot be obtained without too many simplifying approximations so, instead, an accurate value is obtained through simulations.

Figure 5.4 shows the variance of the achievable rate per user as a function of the cluster size M. The different curves are for different system parameters, for instance, several values of the SNR and different antenna configurations. Other simulation parameters are fixed, as is the path loss exponent $\gamma = 3.8$, which is a rather typical value for urban environments as used in [14], and the uniform power assignment that is considered.

It is interesting to note that increasing the degrees of freedom available in the system, by using a higher number of antennas, increases the variance of the achievable rate. Something similar happens when the SNR of the system increases.

An important detail that can be stressed is how the variance does not decrease unboundedly with the cluster size, but a minimum can be found, analogously to the maximum that can be found in the mean achievable rate in Subsection 4.5.3.

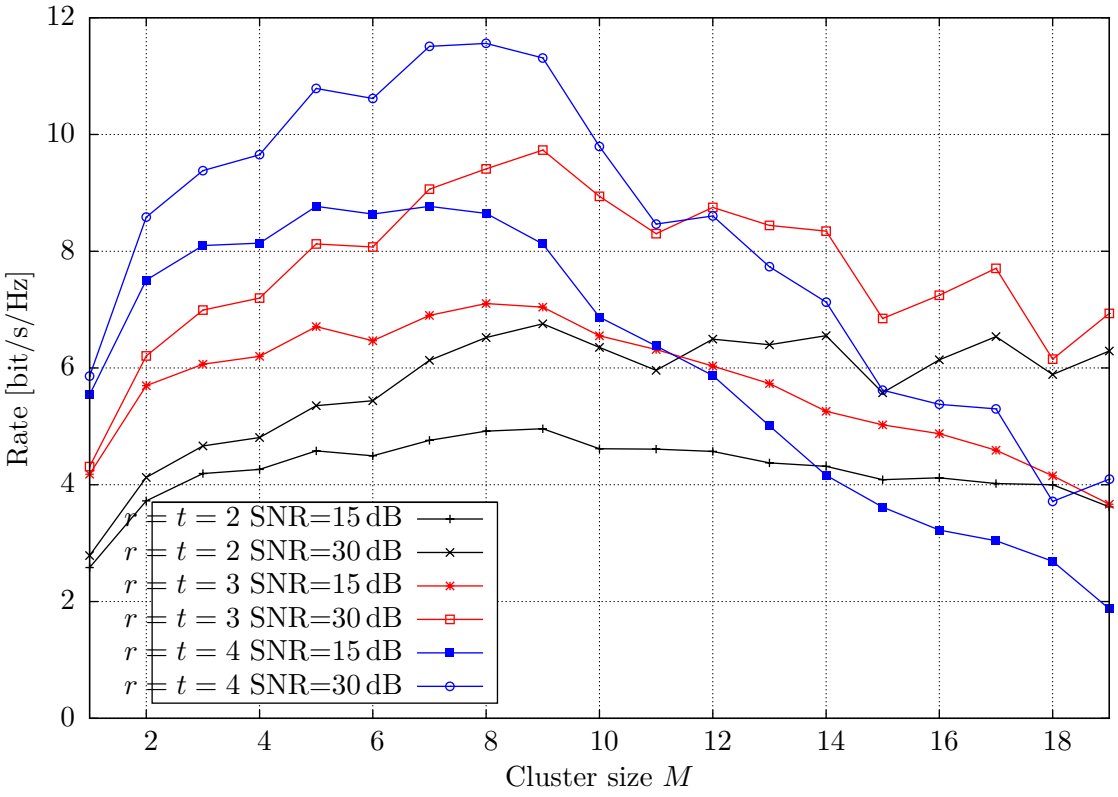


Figure 5.5: Rate achieved by 90 % of the users with different antenna configurations and different values of SNR.

Understanding that fairness in the system is closely related to the variance of the achievable rates per user, it is plain to see that in order to maximize the fairness of the system, a limited number of BSs per cluster is preferable.

5.3 Fairness and QoS considerations

Using the statistics of the coordination schemes considered that have been presented in Section 5.2, several fairness and QoS characteristics can be studied.

One of these is the minimum rate that can be guaranteed to a percentage X of the users. This is given by the value of the rate, the abscise, when the CDF is equal to $(1 - {}^{X}/_{100})$, which gives the rate that can be guaranteed to X % of the users.

Figure 5.5 presents the rate achieved by 90 % of the users as a function of the cluster size M, for different system parameters, considering a uniform power assignment.

The behavior observed suggests, again, that a limited cluster size is advisable in order to deliver the advantages of MIMO to the maximum number of users possible.

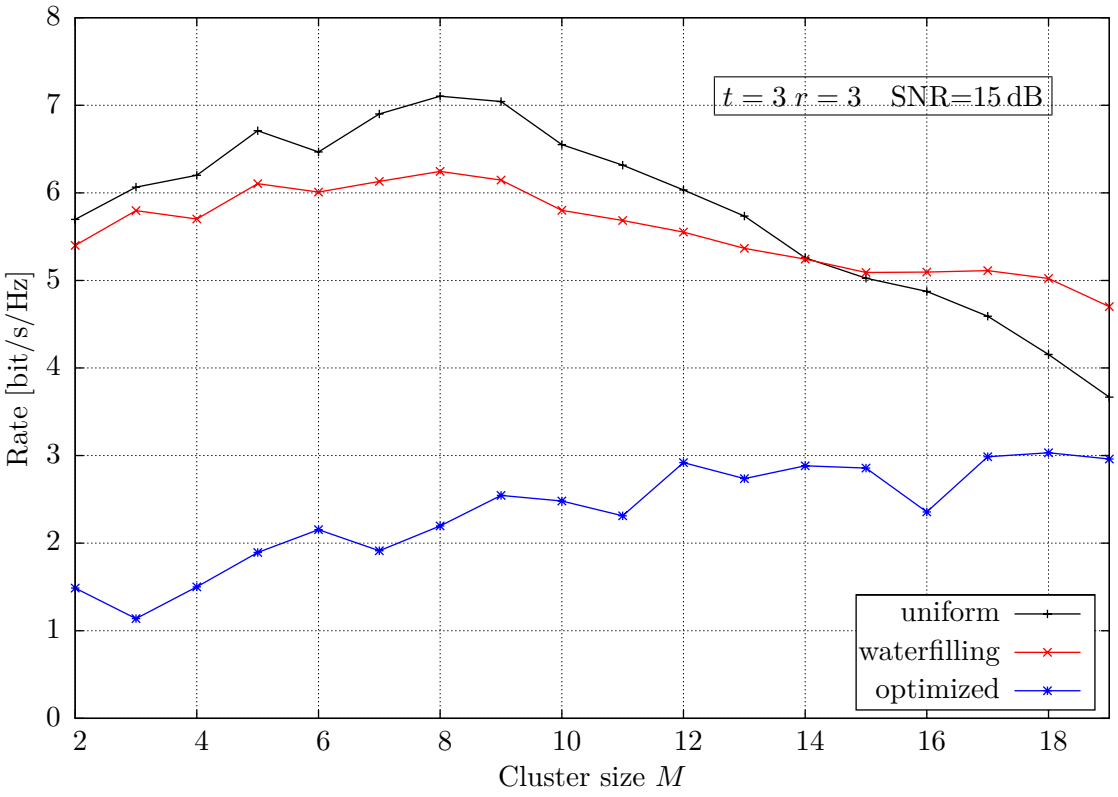


Figure 5.6: Rate achieved by 90 % of the users with different power assignment schemes, $t = r = 3$ and SNR= 15 dB.

A maximum of the minimum rate guaranteed to 90 % of the users appears for a cluster size of around 7, and it decreases as the cluster grows bigger. This decrease is steeper for higher number of antennas because the variance increases, as seen in Figure 5.4. In particular, it can be seen that for clusters of size 18, a configuration with $t = r = 4$ performs worse than a configuration with $t = r = 3$.

Note, however, that the power assignment method plays an important role, as it can be seen from the slope of the CDF in Figure 5.3. In order to illustrate this, Figure 5.6 represents the minimum rate that can be guaranteed to 90 % of the users for different power allocation strategies.

The CVX power allocation scheme presents the worst performance, in terms of the minimum guaranteed rate. It is easy to understand that the CVX solution tries to maximize the sum-rate, which comes at the cost of letting the users with the worst conditions out, yielding a very unfair system.

On the other hand, both waterfilling and the uniform assignment schemes perform better in this regard, and they show a similar behavior. The only difference in the behavior appears for bigger clusters. The uniform power assignment performs better for small clusters, and the minimum rate decreases noticeably for large clusters. The modified water-filling alterna-

tive shows a much more constant behavior, with much less difference between small and big clusters.

Chapter 6

Adaptive Block Diagonalization and User Scheduling With Out of Cluster Interference*

6.1 Introduction

In Chapter 4 and 5 the performance of a clustered network using BD has been analyzed, both in terms of mean achievable rate per user and in terms of network fairness. In both studies, a common result appears: a limited cluster size can be beneficial. This is true not only regarding the rate that the users can achieve, but also considering the requirements imposed on the network infrastructure. When the size of the cellular network grows, global coordination becomes impractical, due to the increased feedback and backhaul requirements. Additionally, there are theoretical works that show how the gains from coordination are intrinsically limited for an increasing network size [10].

The main drawback of clustering is the presence of OCI, and in particular BD performs poorly when OCI is considered [38]. The problem approached in this chapter is the performance loss of BD when OCI is present. A simple and practical algorithm is presented, based on a hybrid strategy combining BD and Single User (SU) processing. The best transmission strategy is chosen according to a metric that is compared with a simple threshold at each user equipment.

The scenario here considered is a multiuser network, with each cell serving multiple users. A low-complexity algorithm is proposed to schedule the users, trying to take advantage of the multiuser diversity to increase the mean rate per user. In [39] a similar suboptimal algorithm, based on the Frobenius norm of the channel matrix is proposed, but it is not analyzed in the presence of OCI nor is it combined with a hybrid precoding strategy.

*The work shown in this chapter has been presented in [37].

6.2 System Model

The system model used in this chapter is the same as the clustered network model in Chapter 4. It is focused on the downlink of a cellular network with a set of $\mathcal{B} = \mathcal{B}_{\text{in}} \cup \mathcal{B}_{\text{out}}$ of cells, where $\mathcal{B}_{\text{in}}$ is the set of M cells that form the cluster under study, $\mathcal{B}_{\text{out}}$ represents the set of M_{interf} cells external to the cluster, and $\mathcal{B}_{\text{in}} \cap \mathcal{B}_{\text{out}} = \emptyset$. Again, it will be considered that each user is associated to one and only one BS.

The signal received at the i-th user equipment is then given by

$$y_i = \boldsymbol{H}_i \boldsymbol{W}_i^{(\text{tx})} \boldsymbol{s}_i + \underbrace{\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\text{tx})} \boldsymbol{s}_j}_{\text{Inner Interference}} + \underbrace{\sum_{k \in \mathcal{B}_{\text{out}}} \overline{\boldsymbol{H}}_{ik} \overline{\boldsymbol{x}}_k}_{\text{OCI}} + \boldsymbol{n}_i \tag{6.1}$$

which is another way of writing (3.18) when external interference is considered. $\overline{\boldsymbol{H}}_{ik} \in \mathbb{C}^{r \times t}$ is the channel matrix from the k-th BS outside the cluster to the i-th user in the cluster, and $\overline{\boldsymbol{x}}_k$ is the transmitted signal at the k-th BS outside the cluster.

Analogously to (3.19), the interference and noise terms in (6.1) can be grouped together

$$\hat{\boldsymbol{z}}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\text{tx})} \boldsymbol{s}_j + \sum_{k \in \mathcal{B}_{\text{out}}} \overline{\boldsymbol{H}}_{ik} \overline{\boldsymbol{x}}_k + \boldsymbol{n}_i. \tag{6.2}$$

The ergodic rate obtained at the i-th receiver can then be written as

$$R_i = \log_2 \left| \boldsymbol{I} + \boldsymbol{H}_i \boldsymbol{W}_i^{(\text{tx})} \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_i^{(\text{tx}),H} \boldsymbol{H}_i^H \boldsymbol{R}_{\hat{\boldsymbol{z}}_i}^{-1} \right| \tag{6.3}$$

where

$$\begin{aligned} \boldsymbol{R}_{\hat{\boldsymbol{z}}_i} &= \mathbb{E}\left\{ \hat{\boldsymbol{z}}_i \hat{\boldsymbol{z}}_i^H \right\} \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\text{tx})} \boldsymbol{R}_{\boldsymbol{s}_j} \boldsymbol{W}_j^{(\text{tx}),H} \boldsymbol{H}_i^H + \sum_{k \in \mathcal{B}_{\text{out}}} \overline{\boldsymbol{H}}_{ik} \boldsymbol{R}_{\overline{\boldsymbol{x}}_k} \overline{\boldsymbol{H}}_{ik}^H + \sigma^2 \boldsymbol{I} \end{aligned} \tag{6.4}$$

is the covariance matrix of the interference plus noise vector in (6.2).

The rate expression (6.3) depends on the transmission strategy used within the cells of the cluster, represented by the precoding matrix $\boldsymbol{W}_i^{(\text{tx})}$ for $i \in \mathcal{B}_{\text{in}}$. In the current work two transmission strategies are considered:

- Block Diagonalization.
- Single User Processing.

6.3 Transmission Strategy

6.3.1 Block Diagonalization

BD transmission strategy has been thoroughly described in Section 3.3 in the absence of interference coming from outside the cluster that is coordinated using BD.

In this work, OCI is considered, and in this case BD is not able to remove it, as the BSs only coordinate to get rid of the interference from within the cluster. As a result, the rate obtained when using BD is no longer given by (3.39), but the following revised expression

$$R_i^{\mathrm{BD}} = \log_2 \left| \boldsymbol{I} + \widehat{\boldsymbol{\Lambda}}_i \boldsymbol{R}_{\boldsymbol{s}_i} \widehat{\boldsymbol{\Lambda}}_i^H \left(\sum_{k \in \mathcal{B}_{\mathrm{out}}} \overline{\boldsymbol{H}}_{ik} \boldsymbol{R}_{\overline{\boldsymbol{x}}_k} \overline{\boldsymbol{H}}_{ik}^H + \sigma^2 \boldsymbol{I} \right)^{-1} \right| \tag{6.5}$$

where the additional term due to the OCI is present and $\boldsymbol{R}_{\overline{\boldsymbol{x}}_k} = \mathbb{E}\left\{ \overline{\boldsymbol{x}}_k \overline{\boldsymbol{x}}_k^H \right\} \in \mathbb{C}^{t \times t}$ is the covariance matrix of the signal transmitted from the k-th BS outside the cluster.

6.3.2 Single User Processing

The other transmission strategy that is considered in this work is Single User processing. This is the case when no coordination is used, so that each BS serves only one user and, therefore, all the rest of BSs are considered interferers.

This translates into

$$\boldsymbol{W}_{ij}^{(\mathrm{tx})} = \boldsymbol{0} \forall j \neq i. \tag{6.6}$$

The signal received at the i-th user can be written as

$$y_i = \boldsymbol{H}_{ii} \boldsymbol{x}_i + \underbrace{\sum_{\ell \in \mathcal{B}_{\mathrm{in}} \setminus \{i\}} \boldsymbol{H}_{i\ell} \boldsymbol{x}_\ell}_{\text{Inner Interference}} + \underbrace{\sum_{k \in \mathcal{B}_{\mathrm{out}}} \overline{\boldsymbol{H}}_{ik} \overline{\boldsymbol{x}}_k}_{OCI} + \boldsymbol{n}_i \tag{6.7}$$

Provided (6.7), if the SVD of the channel matrix $\boldsymbol{H}_{ii}$ is considered

$$\boldsymbol{H}_{ii} = \boldsymbol{U}_i \Lambda_i \boldsymbol{V}_i^H \tag{6.8}$$

then the following precoding matrix can be used in order to maximize the rate of the i-th user [40]

$$\boldsymbol{W}_{ii}^{(\mathrm{tx})} = \boldsymbol{V}_i \tag{6.9}$$

so that, using $\boldsymbol{U}_i^H$ as the receive filter, the achievable rate becomes

$$R_i^{\mathrm{SU}} = \log_2 \left| \boldsymbol{I} + \boldsymbol{\Lambda}_i \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{\Lambda}_i^H \left(\sum_{\ell \in \mathcal{B}_{\mathrm{in}} \setminus \{i\}} \boldsymbol{H}_{i\ell} \boldsymbol{R}_{\boldsymbol{x}_\ell} \boldsymbol{H}_{i\ell}^H + \sum_{k \in \mathcal{B}_{\mathrm{out}}} \overline{\boldsymbol{H}}_{ik} \boldsymbol{R}_{\overline{\boldsymbol{x}}_k} \overline{\boldsymbol{H}}_{ik}^H + \sigma^2 \boldsymbol{I} \right)^{-1} \right| \tag{6.10}$$

6.3.3 Transmission Strategy Selection

In [41] it is shown that the maximum capacity of a Multiple Input Single Output (MISO) downlink channel can be reached using a combination of two transmission strategies, the optimal Maximum Ratio Combining (MRT) and Zero Forcing Beamforming (ZFBF). Based on this, [42] presents a method for the users to decide locally and individually the most convenient transmission strategy from the two options presented in [41]. The way to do this is through a threshold on the SINR, and a closed form expression for this threshold is provided in [42].

In [43], a similar result to that of [41] is presented for the MIMO case, but the solution offered, apart from not having a closed form expression, was based on sequentially solving a series of optimization problems, which renders its applicability rather difficult.

In the current work, the same approach as in [42] is followed. The proposal is to use a metric that can be easily calculated at each receiver, and to compare this metric with a fixed threshold locally at each user in order to decide which transmission strategy to use.

Intuitively, BD will perform better when the OCI is low, compared to the power received from the BSs in the cluster. Hence, the proposed metric is

$$\theta_i = \frac{\sum_{j \in \mathcal{B}_{\mathrm{in}}} \mathrm{Tr}\left(\boldsymbol{H}_{ij} \boldsymbol{R}_{\boldsymbol{x}_j} \boldsymbol{H}_{ij}^H \right)}{\sum_{k \in \mathcal{B}_{\mathrm{out}}} \mathrm{Tr}\left(\overline{\boldsymbol{H}}_{ik} \boldsymbol{R}_{\overline{\boldsymbol{x}}_k} \overline{\boldsymbol{H}}_{ik}^H \right)} \tag{6.11}$$

which is the quotient of the power received from the BSs in the cluster and the power received from the BSs outside the cluster.

After the metric is computed, it is compared with a fixed threshold θ_{th} for which there is no closed form expression. That is the reason why the threshold considered in this work is calculated through simulations and it is assumed to be known by all the users. The details of how this threshold is calculated are given in Section 6.6.

The decision about what transmission strategy to use is made locally by each user. The i-th user compares the metric θ_i with the threshold θ_{th}, and decides to use BD if $\theta_i > \theta_{\text{th}}$, and choosen SU if $\theta_i \leq \theta_{\text{th}}$.

All the users feed back their decissions to their BS, and this information is jointly used by the BSs in the cluster to coordinate the scheduling of the users and the transmission strategy used for each of them.

6.4 Scheduling

After the users have made their decision and fed it back to the BSs, these will know which users are more suited to being served using BD and which ones using SU.

Analogously to [42], the users are grouped so that the transmission strategy in all the BSs is the same within a given transmission interval. The motive for this was simplicity, in [42], whereas in this work it is proposed to guarantee a good performance. This is so because users served with BD will experience a serious degradation in their rate if not all the BSs in the cluster coordinate, i.e., some of the BSs transmit to their users using SU precoding.

Users that are better served using SU are indiferent to other users' strategies, as no power control is considered or used, and all BSs will be transmitting at maximum power. On the other hand, when the transmission strategy used is BD, which users are selected in each cell is an important operating decision. Depending on the channel matrices of each user, the BD process will result in a higher or a lower rate. The objective is, then, to group the users from different cells so that a certain metric is maximized. In particular, the metric used in this work will be related to the achievable sum-rate.

In [44] a similar approach is proposed for a MISO scenario, where users are scheduled for simultaneous transmission when their channel vectors are as orthogonal as possible. In the MIMO case the channels are not vectors but matrices, and the concept of orthogonal channels is not as clear as in [44]. They propose, nonetheless, an extesion to their user selection algorithm that can deal with multiantenna users, but it is not applicable here because of the selection of BD as precoder, instead of ZFBF as transmission strategy.

The objective of the current work is to group BD users so that the result of the BD precoding yields the maximum achievable sum-rate. Equivalently, the users will be scheduled for transmission in groups of M, i.e., one per cell, so that the values of the diagonal of $\widehat{\boldsymbol{\Lambda}}_i$ in (6.5) are maximized.

Given a square matrix $\boldsymbol{A}$, the following identity relating its trace with the sum of its eigenvalues holds

$$\text{Tr}\left(\boldsymbol{A}\right) = \text{Tr}\left(\text{eig}\left(\boldsymbol{A}\right)\right) \tag{6.12}$$

where $\text{eig}\left(\boldsymbol{A}\right)$ is a diagonal matrix whose elements are the eigenvalues of the matrix $\boldsymbol{A}$.

Algorithm 1 BD User Selection

Sort the set $\mathcal{U}_{\text{BD}}$ in decreasing order of θ_i
while $|\mathcal{U}_{\text{BD}}| \geq M$ **do**
  $i = \text{first}\,(\mathcal{U}_{\text{BD}})$
  $\mathcal{U}_g = \{i\}$
  $\mathcal{U}_{\text{BD}} = \mathcal{U}_{\text{BD}} \backslash \{i\}$
  $\boldsymbol{H}_g = \boldsymbol{H}_i$
  $N_g = 1$
  while $N_g \leq M$ **do**
    $j = \underset{j \in \mathcal{U}_{\text{BD}} \backslash\, \text{cell}(\mathcal{U}_g)}{\text{arg. max.}} \left\| \begin{bmatrix} \boldsymbol{H}_g \\ \boldsymbol{H}_j \end{bmatrix} \right\|_F$
    $H_g = \begin{bmatrix} \boldsymbol{H}_g \\ \boldsymbol{H}_j \end{bmatrix}$
    $N_g = N_g + 1$
    $\mathcal{U}_g = \mathcal{U}_g \cup \{j\}$
    $\mathcal{U}_{\text{BD}} = \mathcal{U}_{\text{BD}} \backslash \{i\}$
  end while
end while

Additionally, the Frobenius norm of a matrix $\boldsymbol{A}$ is defined as

$$\|\boldsymbol{A}\|_F = \sqrt{\text{Tr}\,(\boldsymbol{A}\boldsymbol{A}^H)} \tag{6.13}$$

Combining (6.12) and (6.13), the following relation can be expressed

$$\text{Tr}\,(\text{eig}\,(\boldsymbol{A}\boldsymbol{A}^H)) = \|\boldsymbol{A}\|_F^2 \tag{6.14}$$

can be used to have a measure of the magnitude of the eigenvalues of the matrix $\boldsymbol{A}$, and it is a good candidate as metric to group the users for BD, similarly to what is done in [39].

Given the i-th user's channel matrix $\boldsymbol{H}_i$, in order to search for the user j that will yield the maximum sum-rate using BD, the idea is to look for the user j with channel matrix $\boldsymbol{H}_j$ that maximizes the Frobenius norm of the compound matrix, because

$$\text{Tr}\left(\widehat{\boldsymbol{\Lambda}}_{ij}\right) \propto \left\| \begin{bmatrix} \boldsymbol{H}_i \\ \boldsymbol{H}_j \end{bmatrix} \right\|_F^2 \tag{6.15}$$

Algorithm 1 is proposed to form the groups of M users that maximize the rate using BD. First the set of users that want to be served using BD, $\mathcal{U}_{\text{BD}}$, is sorted in descending order, with respect to the magnitude of the metric (6.11). Then the groups of M users are generated by adding one user at a time, using the Frobenius norm of the resulting matrix as a measure of the magnitude of the singular values after performing the BD, as suggested by (6.15).

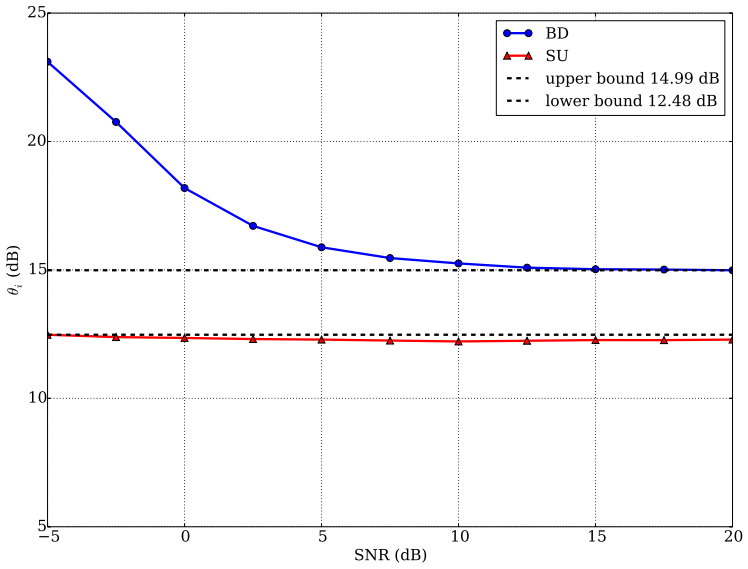


Figure 6.1: Mean value of the metric for different SNR values, in the presence of OCI for a 7 cell cluster with 2x2 antennas configuration.

In Algorithm 1, the functions *first* and *cell* refer to getting the first element in an ordered set, and to returning the set of users in the same cells as the users in the argument set, respectively. A possible result of Algorithm 1 is that not all the users that have selected BD as their preferred transmission strategy can be fit in a group with other BD users. In that case, the excess users are served using SU.

Finally, a round-robin strategy is used to transmit to all the groups that have been formed, both the BD and SU groups.

6.5 Transmission strategy threshold computation

In Section 6.4 the threshold θ_{th} was introduced as the means for the users to decide between the two possible transmission strategies. This threshold is assumed to be fixed, and it is precalculated via simulations.

In order to calculate the threshold a single user is placed in each of the cells of the cluster, then the rate is calculated both when all the BS coordinate to transmit using BD and when they transmit independently using SU. Apart from the rate for the two transmission strategies, the metric in (6.11) is also computed. This process is repeated for multiple user locations and channel realizations.

With the data gathered through the simulations described, it is possible to get Figure 6.1 where each of the solid curves represents the mean value of the metric in (6.11), in dB, for the

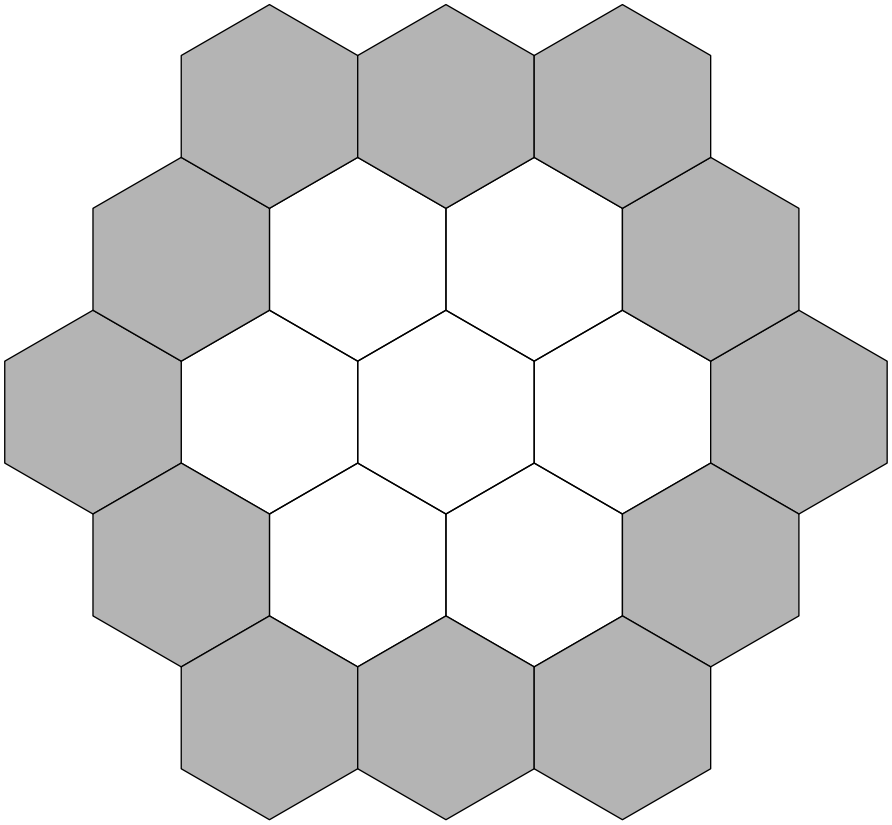


Figure 6.2: The cells in the cluster (white) experience the OCI generated by the interfering cells (shaded).

users whose rate is higher using BD (blue), and for those who are better off being served using SU (red). The dashed curves bound a gap between both curves, that will be used to select the threshold ϕ_{th} for (6.11). For the example in Figure 6.1, the value chosen for the threshold is the mean of the *upper* and *lower* bounds in the figure, yielding a value of $\theta_{\text{th}} = 13.74\,\text{dB}$.

As it has already been said, this value of the threshold is assumed to be fixed and it has to be precomputed for the particular scenario under study. Figure 6.1 is calculated for the scenario represented in Figure 6.2, which is also used in Section 6.6 to asses the performance of the algorithm proposed in this work.

Appendix B offers a characterization of the behavior of θ_{th} for different values of the scenario and system parameters, so that its dependence on these factors can be observed.

6.6 Performance Analysis

The cellular scenario considered in this chapter differs slightly from the scenario used in Chapter 4 and 5. A single cluster of 7 cells is considered, in a hexagonal layout and surrounded by a single tier of cells that will account for the OCI.

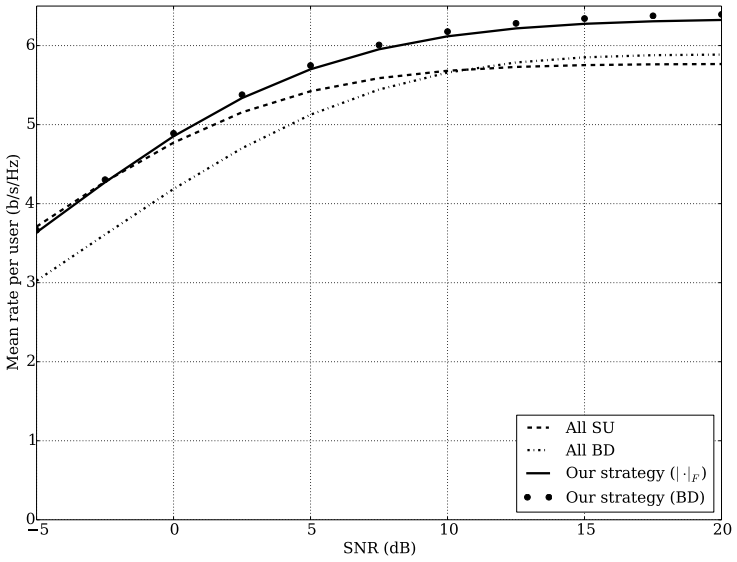


Figure 6.3: Mean rate obtained for a 2x2 scenario in the presence of OCI, 100 users per cell.

Figure 6.2 shows the scenario used in this chapter, where the white cells form the cluster under study, and the shaded cells represent the out of cluster interferers.

At each simulation run, 100 iid users are randomly placed within each cell of the cluster, according to a uniform spatial distribution over each cell. Each of the users has the same antennas as each of the BSs, i.e., $t = r$, either 2 or 3 in the simulations.

The channel model used is described in Section 3.2.

After all the users are placed in the scenario, they are scheduled for transmision and immediately after that the rate is calculated for the following transmission options:

- All BSs transmit using SU.
- All BSs transmit using BD.
- The transmission strategy is chosen using the algorithm and scheduling proposed in this paper.
- The same as the previous, but the scheduling is performed based on the rates obtained using BD instead of the approximation in (6.15).

In all cases, the power assignment is done using the scaled water-filling described in Subsection 3.4.3, in order to accomodate the PBPC.

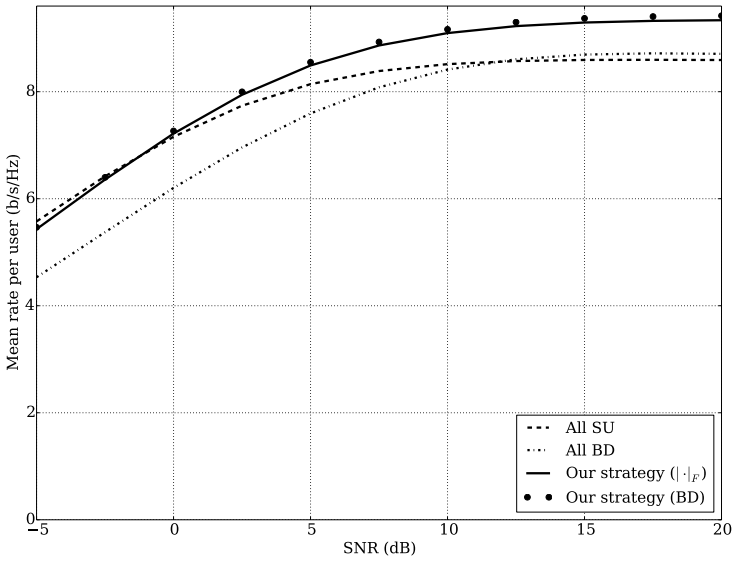


Figure 6.4: Mean rate obtained for a 3x3 scenario in the presence of OCI, 100 users per cell.

Figure 6.3 and Figure 6.4 show the mean achievable rate per user for the transmission options considered, and for two MIMO configurations, $t = r = 2$ and $t = r = 3$ respectively. First thing that can be seen is how the OCI severely degrades the performance of BD, especially in the low SNR regime, where it actually performs worse than not using coordination at all. The user of the mixed strategy proposed in this work improves the performance over the whole SNR range. It is important to note how the approximation suggested in (6.15) is highly accurate, compared to using BD to calculate the rates for the scheduling algorithm. And not only is it accurate, but it also is much simpler to implement and much less computationally expensive [†]

In Figure 6.5 the mean achievable rate per user is presented as a function of the number of users per cell, for a fixed value of SNR of 10 dB. The improvement introduced by the proposed scheme with respect to both BD and SU increases with the number of users per cell. This is easy to explain, as the more users there are in each cell, the increased multiuser diversity makes it more likely to find the right users to form a group.

It is clear, from the previous results, that the OCI has a very serious impact on the performance of BD, but this is even more severe when the fairness of the rates of all users is considered. Figure 6.6 represents the CDF of the rates obtained with each of the transmission strategies for a fixed value of SNR of 10 dB. It can be seen how the rates obtained using the mixed strategy are always higher than using each of the strategies, BD or SU, independently. This difference in favor of the mixed strategy is even higher for the users with the lowest rates,

[†]As described in Chapter 3 the BD computation involves two SVD, which has an overall computational complexity of $O(mn^2)$ for a matrix of dimmensions $m \times n$, [45]. In contrast the computation of Frobenius norm, which essentially requires the addition of all the elements of a matrix, with a complexity of $O(mn)$.

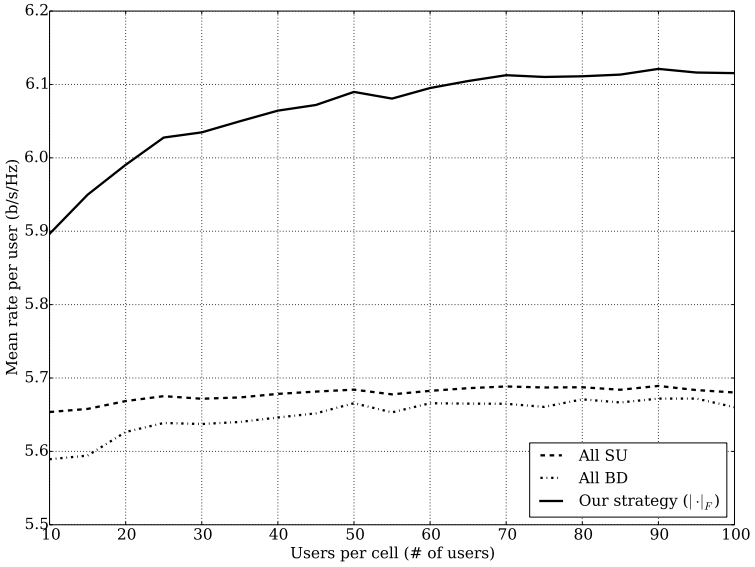


Figure 6.5: Mean rate for a 2x2 scenario as a function of the number of users per cell, for an SNR of 10 dB.

which indicates an improvement in the fairness of the system.

Figure 6.7 shows the average rate for the 5 % worst users. For these users, BD in the presence of OCI performs poorly, and the use of the hybrid strategy proposed in this work allows to recover from the loss due to the OCI, and to match the performance obtained using SU.

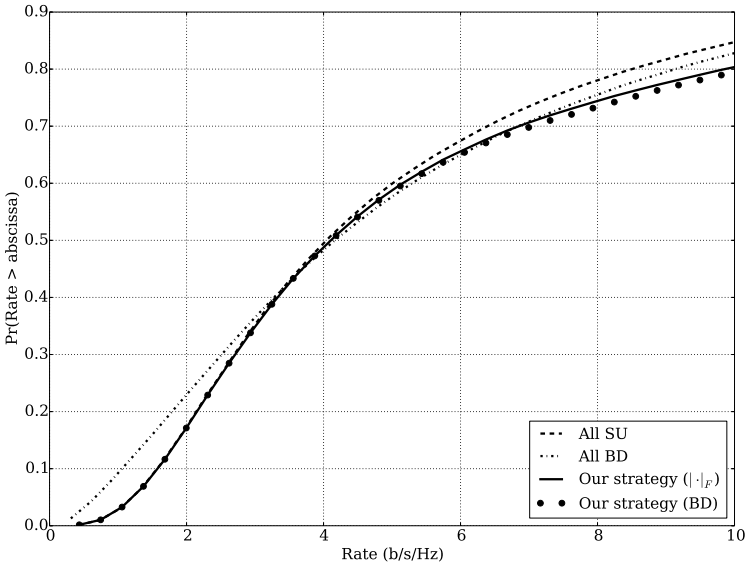


Figure 6.6: CDF of the rates obtained in a 2x2 scenario, with an SNR of 10 dB.

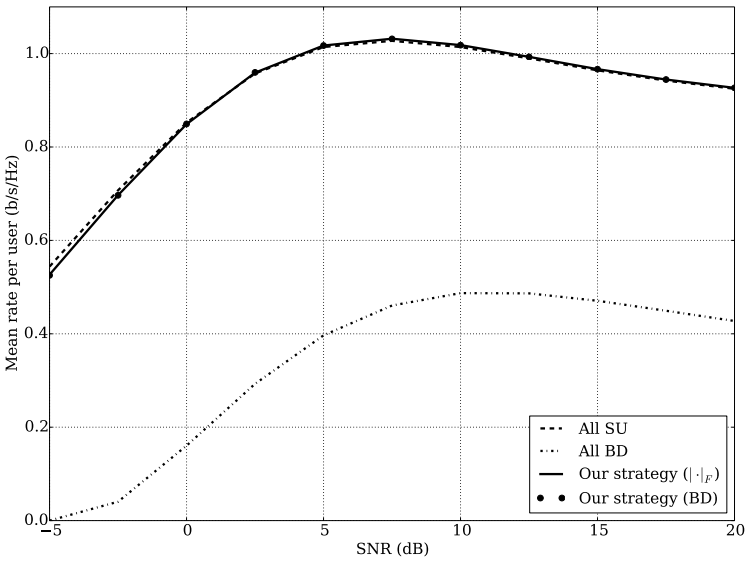


Figure 6.7: Mean rate of the 5% worst users, in a 2x2 scenario in the presence of OCI, 100 users per cell.

Chapter 7

Conclusions

Appendix A

Derivation of Z_i

Using the change of variable

$$v = \frac{u}{R_{\text{cell}}} \tag{A.1}$$

in (4.38), it becomes

$$Z_i = \int_0^1 \log\left(\frac{v^{-\gamma}}{\frac{\sigma_i^2}{P_{\max}R_{\text{cell}}^{-\gamma}} + M_{\text{eq},i}\left(\bar{d}_i - v\right)^{\gamma}}\right) 2v\mathrm{d}v \tag{A.2}$$

Recall the definition of the SNR ρ in (3.25), then

$$\begin{aligned} Z_i &= \int_0^1 \log\left(\frac{v^{-\gamma}}{\frac{1}{\rho} + M_{\text{eq},i}\left(\bar{d}_i - v\right)^{-\gamma}}\right) 2v\mathrm{d}v \\ &= 2\underbrace{\int_0^1 \log\left(\frac{1}{\frac{1}{\rho} + M_{\text{eq},i}\left(\bar{d}_i - v\right)^{-\gamma}}\right) v\mathrm{d}v}_{\texttt{I}} - 2\gamma\underbrace{\int_0^1 \log\left(v\right) v\mathrm{d}v}_{\texttt{II}} \end{aligned} \tag{A.3}$$

The term `II` in (A.3) is evaluated using (2.723) from [29], namely

$$\int v\log\left(v\right)\mathrm{d}v = v^2\left(\frac{\log\left(v\right)}{2} - \frac{1}{4}\right) \tag{A.4}$$

Using the change of variable

$$w = \bar{d}_i - v \tag{A.5}$$

in the term I of (A.3), it becomes

$$\begin{aligned}(\mathtt{I}) &= -\int_{\bar{d}_i}^{\bar{d}_i-1} \log\left(\frac{\rho}{1+M_{\mathrm{eq},i}\rho w^{-\gamma}}\right)\left(\bar{d}_i - w\right)\mathrm{d}w \\ &= \underbrace{\int_{\bar{d}_i-1}^{\bar{d}_i} \log\left(\rho\right)\left(\bar{d}_i - w\right)\mathrm{d}w}_{\mathtt{III}} + \underbrace{\int_{\bar{d}_i-1}^{\bar{d}_i} \log\left(\frac{1}{1+M_{\mathrm{eq},i}\rho w^{-\gamma}}\right)\left(\bar{d}_i - w\right)\mathrm{d}w}_{\mathtt{IV}}\end{aligned} \tag{A.6}$$

The term III of (A.6) is straightforward to calculate

$$\int_{\bar{d}_i-1}^{\bar{d}_i} \log\left(\rho\right)\left(\bar{d}_i - w\right)\mathrm{d}w = \log\left(\rho\right) \tag{A.7}$$

On the other hand, the term IV in (A.4) can be computed using the following result from [46]

$$\begin{aligned}\int_{\bar{d}_i-1}^{\bar{d}_i} \log\left(\frac{1}{1+M_{\mathrm{eq},i}\rho w^{-\gamma}}\right)\left(\bar{d}_i - w\right)\mathrm{d}w = w\Bigg[&\frac{2\bar{d}_i - w}{2}\log\left(\frac{w^{\gamma}}{M_{\mathrm{eq},i}\rho + w^{\gamma}}\right) \\ &- \bar{d}_i\gamma\,{}_2\mathrm{F}_1\left(1, \frac{1}{\gamma}; \frac{\gamma+1}{\gamma}; \frac{-w^{\gamma}}{M_{\mathrm{eq},i}\rho}\right) \\ &+ \frac{w\gamma}{4}\,{}_2\mathrm{F}_1\left(1, \frac{2}{\gamma}; \frac{\gamma+2}{\gamma}; \frac{-w^{\gamma}}{M_{\mathrm{eq},i}\rho}\right)\Bigg]\end{aligned} \tag{A.8}$$

where

$${}_2\mathrm{F}_1\left(a, b; c; z\right) = \frac{\Gamma\left(c\right)}{\Gamma\left(b\right)\Gamma\left(c-b\right)}\int_0^1 \frac{x^{b-1}\left(1-x\right)^{c-b-1}}{\left(1-x\right)^{a}}\mathrm{d}x \tag{A.9}$$

is the hypergeometric function.

Combining the previous, Z_i can be expressed as

$$
\begin{aligned}
Z_i = & \frac{\gamma}{2} + \log(\rho) + \bar{d}_i^2 \log\left(\frac{\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}}\right) \\
& - 2\bar{d}_i^2 \gamma \, {}_2\mathrm{F}_1\left(1, \frac{1}{\gamma}; \frac{\gamma+1}{\gamma}; \frac{-\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}}\right) \\
& + \frac{\bar{d}_i^2}{2} \gamma \, {}_2\mathrm{F}_1\left(1, \frac{2}{\gamma}; \frac{\gamma+2}{\gamma}; \frac{-\bar{d}_i^{\gamma}}{M_{\mathrm{eq},i}\rho + \bar{d}_i^{\gamma}}\right) \\
& - \left(\bar{d}_i^2 - 1\right) \log\left(\frac{\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho + \left(\bar{d}_i - 1\right)^{\gamma}}\right) \\
& + 2\bar{d}_i \left(\bar{d}_i - 1\right) \gamma \, {}_2\mathrm{F}_1\left(1, \frac{1}{\gamma}; \frac{\gamma+1}{\gamma}; \frac{-\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho}\right) \\
& - \frac{\left(\bar{d}_i - 1\right)^2}{2} \gamma \, {}_2\mathrm{F}_1\left(1, \frac{2}{\gamma}; \frac{\gamma+2}{\gamma}; \frac{-\left(\bar{d}_i - 1\right)^{\gamma}}{M_{\mathrm{eq},i}\rho}\right)
\end{aligned}
\tag{A.10}
$$

Appendix B

Characterization of $\theta_{\mathbf{th}}$

In this chapter, a brief analysis of the threshold θ_{th} is offered in order to show its dependence on some of the scenario parameters, for instance:

- Number of antennas of the MIMO configuration.
- SNR of the system.
- Cluster size.
- Path loss exponent γ.

Figures B.1 through B.4 show this dependence for a series of combinations of parameters.

First, Figure B.3 and Figure B.4 show how the threshold is consistently independent from the SNR, except for very low values of it.

For increasing values of the path loss exponent, the threshold increases as well, as it can be seen in Figure B.1 and Figure B.2. A higher path loss exponent translates into a lower level of interference between adjacent cells, in which case coordination may not help at all, so each BS better serves its own users independently. This is the reason for a higher threshold, which implies that less users will select BD as their preferred transmission strategy.

Another interesting characteristic that is clear in Figure B.2 and Figure B.4 is how the number of antennas plays no role in the threshold.

These two suggestions mean the advantages that MIMO has to offer have no influence on the value of the θ_{th}, and it should depend mainly on the propagation characteristics of the channel.

This claim can be supported by Figure B.1 and Figure B.3, which show the behavior of the threshold with the cluster size, and by the already mentioned dependence on the path loss exponent. Cluster size is represented by the number of hexagonal tiers that form the cluster,

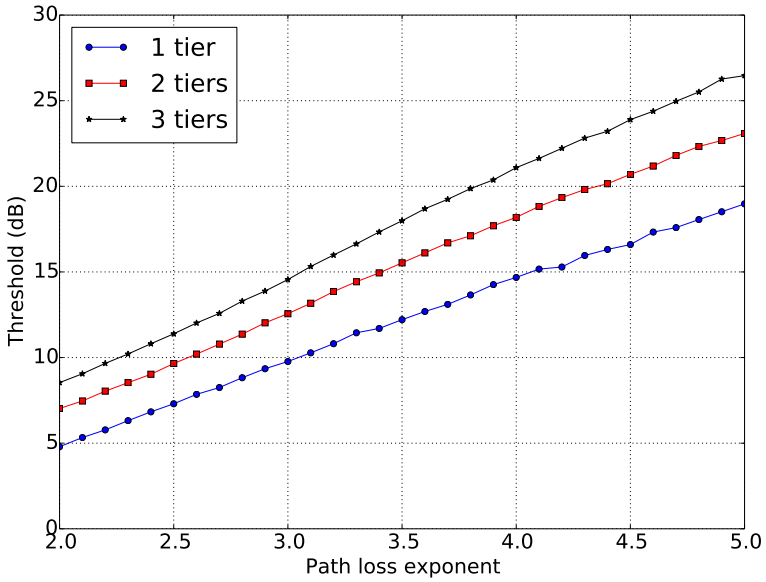


Figure B.1: Threshold as a function of the path loss exponent, for different cluster sizes, for an SNR of 15 dB.

so that 1 tiers is a 7 cells cluster, 2 tiers is a 19 cells cluster, and 3 tiers corresponds to a 37 cells cluster.

Similarly to what happened with the path loss exponent, a bigger cluster means that cells within a cluster may be too far away from some users in the cluster, who may not benefit much from coordination. As observed with the path loss exponent γ, reducing the level of influence among the cells in the cluster increases the value of the threshold, in this case when the cluster considered grows in size. Again, this higher value of θ_{th} means that more users will select SU as their transmission strategy.

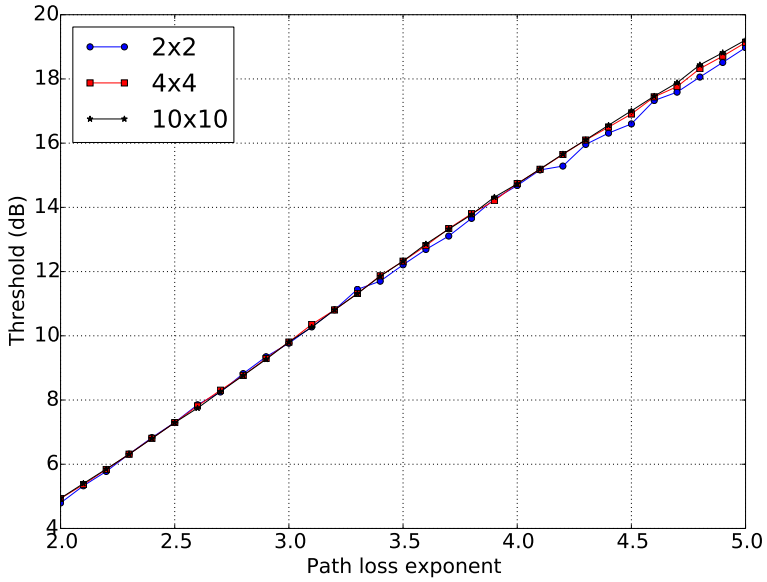


Figure B.2: Threshold as a function of the path loss exponent, for different MIMO configurations, for an SNR of 15 dB.

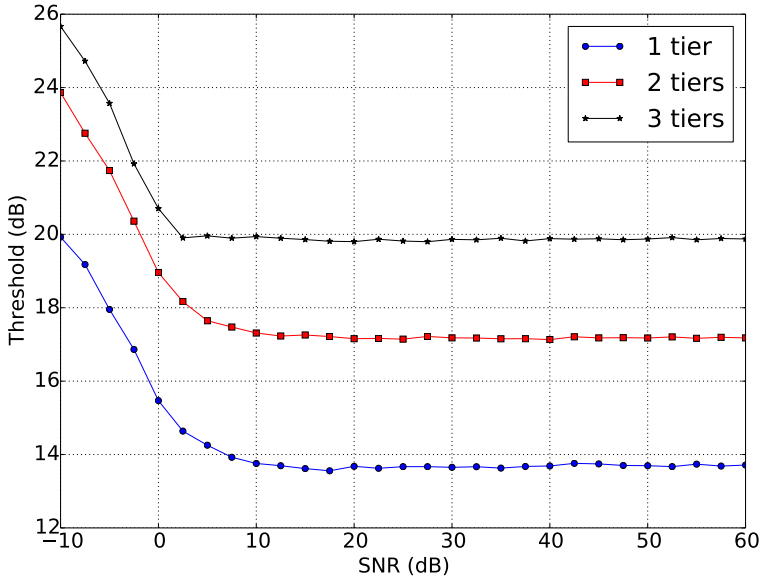


Figure B.3: Threshold as a function of the SNR, for different cluster sizes, for $\gamma = 3.8$.

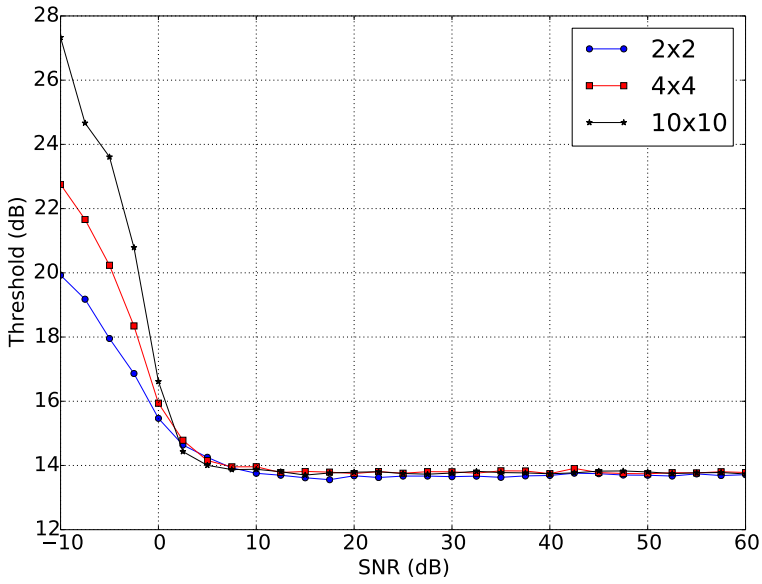


Figure B.4: Threshold as a function of the SNR, for different MIMO configurations, for $\gamma = 3.8$.

Bibliography

[1] 3GPP, "Multiple input multiple output in UTRA (3GPP TR 25.876 version 7.0.0 release 7)," 3GPP, Tech. Rep., Mar. 2007.

[2] a3GPP. (Apr. 2014). LTE, [Online]. Available: `http://www.3gpp.org/LTE`.

[3] ITU-R, "Detailed specifications of the terrestrial radio interfaces of international mobile telecommunications-advanced (IMT-advanced), recommendation ITU-R M.2012-1," ITU-R, Tech. Rep., Feb. 2014.

[4] 3GPP. (Apr. 2014). LTE-Advanced, [Online]. Available: `http://www.3gpp.org/LTE-Advanced`.

[5] ——, "3rd generation partnership project; technical specification group radio access network; coordinated multi-point operation for LTE physical layer aspects (release 11)," 3GPP, Tech. Rep., Dec. 2011.

[6] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *arXiv preprint arXiv:1405.2957*, 2014.

[7] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *Wireless Communications, IEEE*, vol. 21, no. 2, pp. 106–112, Apr. 2014.

[8] D. Gesbert, S. Hanly, H. Huang, S. Shamai (Shitz), O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380–1400, Dec. 2010.

[9] D. Gesbert, S. G. Kiani, A. Gjendemsjø, and G. E. Øien, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2393–2409, Dec. 2007.

[10] A. Lozano, R. W. Heath Jr., and J. G. Andrews, "Fundamental limits of cooperation," *Information Theory, IEEE Transactions on*, vol. 59, no. 9, pp. 5213–5226, 2013.

[11] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[12] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[13] B. Holter, "On the capacity of the MIMO channel: a tutorial introduction," in *Proc. IEEE Norwegian Symposium on Signal Processing*, 2001, pp. 167–172.

[14] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[15] J. Zhang, R. Chen, J. G. Andrews, A. Ghosh, and R. W. Heath Jr., "Networked MIMO with clustered linear precoding," *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 1910–1921, Apr. 2009.

[16] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[17] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 1, pp. 69–73, Jan. 2007.

[18] S.-R. Lee, S.-H. Moon, J.-S. Kim, and I. Lee, "Capacity analysis of distributed antenna systems in a composite fading channel," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 3, pp. 1076–1086, Mar. 2012.

[19] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2009.

[20] A. García Armada, M. Sánchez-Fernández, and R. Corvaja, "Constrained power allocation schemes for coordinated base station transmission using block diagonalization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, pp. 1–14, 2011.

[21] J. M. Cioffi, *Ee379c - advanced digital communication - class notes - chapter 4*, p. 307.

[22] R. L. Burden and J. D. Faires, *Numerical Analysis*. Cengage Learning, 2010.

[23] R. Corvaja, J. J. García Fernández, and A. García Armada, "Mean achievable rates in clustered coordinated base station transmission with block diagonalization," *Communications, IEEE Transactions on*, vol. 61, no. 8, pp. 3483–3493, Aug. 2013.

[24] E. Björnson, N. Jaldén, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 6086–6101, Dec. 2011.

[25] D. B. Cheikh, J.-M. Kelif, M. Coupechoux, and P. Godlewski, "SIR distribution analysis in cellular networks considering the joint impact of path-loss, shadowing and fast fading," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, pp. 1–10, 2011.

[26] B. Pijcke, M. Zwingelstein-Colin, M. Gazalet, M. Gharbi, and P. Corlay, "An analytical model for the intercell interference power in the downlink of wireless cellular networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2011, no. 1, pp. 1–20, 2011.

[27] A. Papoulis, *The Fourier Integral and its Applications*. Mc Graw-Hill, 1962.

[28] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, Jun. 2004.

[29] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, 2000.

[30] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[31] M. C. Grant and S. P. Boyd, *CVX: matlab software for disciplined convex programming, version 2.1*, `http://cvxr.com/cvx`, Mar. 2014.

[32] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. D. Blondel, S. P. Boyd, and H. Kimura, Eds., `http://stanford.edu/~boyd/graph_dcp.html`, Springer-Verlag Limited, 2008, pp. 95–110.

[33] R. Corvaja, J. J. García Fernández, and A. García Armada, "Achievable rate and fairness in coordinated base station transmission," *Communications Letters, IEEE*, vol. 18, no. 4, pp. 584–587, Apr. 2014.

[34] R. W. Heath Jr., M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using poisson point processes," *Signal Processing, IEEE Transactions on*, vol. 61, no. 16, pp. 4114–4126, Aug. 2013.

[35] S. Al-Ahmadi and H. Yanikomeroglu, "On the approximation of the generalized-k distribution by a gamma distribution for modeling composite fading channels," *Wireless Communications, IEEE Transactions on*, vol. 9, no. 2, pp. 706–713, Feb. 2010.

[36] S. Atapattu, C. Tellambura, and H. Jiang, "A mixture gamma distribution to model the SNR of wireless channels," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 12, pp. 4193–4203, Dec. 2011.

[37] J. J. García Fernández, A. García Armada, J. Rubio, A. Pascual-Iserte, O. Font-Bach, and N. Bartzoudis, "Adaptive block diagonalization and user scheduling with out of cluster interference," in *European Wireless 2014; 20th European Wireless Conference; Proceedings of*, May 2014, pp. 1–6.

[38] S. Shim, J. S. Kwak, R. W. Heath Jr., and J. G. Andrews, "Block diagonalization for multi-user MIMO with other-cell interference," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 7, pp. 2671–2681, 2008.

[39] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath Jr., and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *Signal Processing, IEEE Transactions on*, vol. 54, no. 9, pp. 3658–3663, 2006.

[40] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[41] J. Zhang and J. G. Andrews, "Adaptive spatial intercell interference cancellation in multicell wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1455–1468, 2010.

[42] S.-H. Moon, C. Lee, S.-R. Lee, and I. Lee, "Joint user scheduling and adaptive intercell interference cancelation for MISO downlink cellular systems," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 1, pp. 172–181, Jan. 2013.

[43] Z. K. M. Ho and D. Gesbert, "Balancing egoism and altruism on MIMO interference channel," *CoRR*, vol. abs/0910.1688, 2009.

[44] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[45] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed., ser. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2012, vol. 3.

[46] W. A. LLC. (2014). Wolfram alpha, [Online]. Available: `http://www.wolframalpha.com`.