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.
$\operatorname{rank}(\boldsymbol{A})$	Rank of the matrix $\boldsymbol{A}$.
$\ker(\boldsymbol{A})$	Kernel/null-space of the matrix $\boldsymbol{A}$.
$\operatorname{blkdiag}(\boldsymbol{A}_1, \dots, \boldsymbol{A}_N)$	Block diagonal matrix formed with the matrices $\{\boldsymbol{A}_1, \dots, \boldsymbol{A}_N\}$.
$\operatorname{diag}(a_1, \dots, a_N)$	Diagonal matrix whose main diagonal is $\{a_1, \dots, a_N\}$.
$\left\|\boldsymbol{a}\right\|_2$	Norm-2 (Euclidean norm) of the vector, see (1.49).
$\boldsymbol{\nabla}_{\boldsymbol{x}} f(\boldsymbol{x})$	Gradient of a function $f(\boldsymbol{x})$, (1.53)
$\mathbb{E}\{\cdot\}$	Statistical expectation.
$\log_2(\cdot)$	Base 2 logarithm.
$\ln(\cdot)$	Natural (base e) logarithm.
$[a]^{+}$	$\max(0, a)$.

Acronyms

AWGN Additive White Gaussian Noise.

BD Block Diagonalization.

BS Base Station.

iid independent identically distributed.

KKT Karush-Kuhn-Tucker.

PAPC Per Antenna Power Constraint.

PBPC Per Base Station Power Constraint.

pdf probability density function.

SNR Signal to Noise Ratio.

SVD Singular Value Decomposition.

TPC Total Power Constraint.

Chapter 1

System Model

1.1 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 1.2a) 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 1.2b) , 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 Base Station (BS) that is equipped with t transmit antennas. Each of the users considered in the system has r receive antennas. Figure 1.1 shows a schematic representation of such a network, where $Rcell$ is the cell radius, and d_{ij} is the distance from the j-th BS to the i-th user.

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

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n} \tag{1.1}$$

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

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_N \end{bmatrix} \in \mathbb{C}^{Nr\times 1} \tag{1.2}$$

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

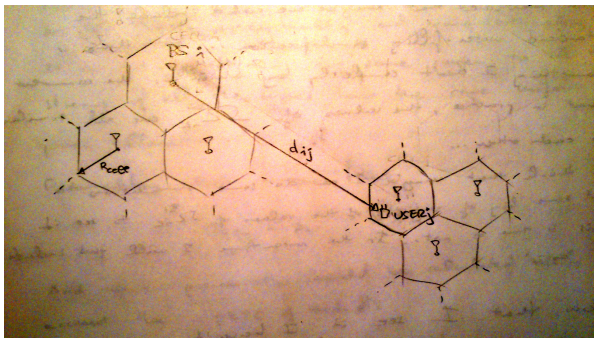
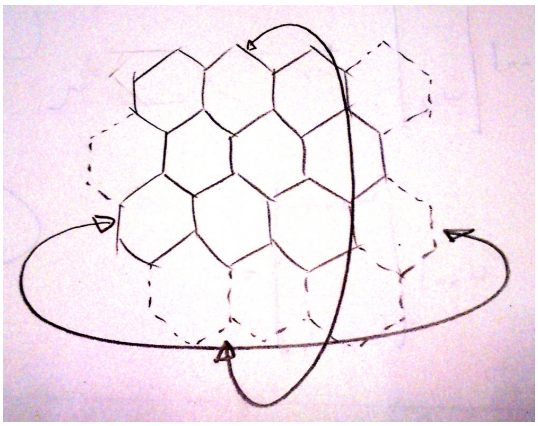
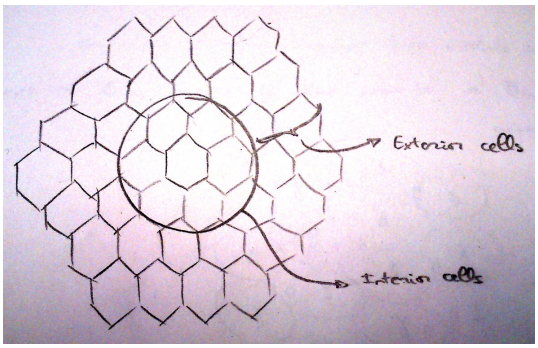


Figure 1.1: Schematic representation of the cellular network that will be used throughout this dissertation.



(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.

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

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{11} & \cdots & \boldsymbol{H}_{1M} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{N1} & \cdots & \boldsymbol{H}_{NM} \end{bmatrix} \in \mathbb{C}^{Nr \times Mt} \tag{1.3}$$

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.

$\boldsymbol{x}$ is the signal transmitted from all the BS, and it is composed of

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_M \end{bmatrix} \in \mathbb{C}^{Mt \times 1} \tag{1.4}$$

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

$$P_{j,\mathrm{tx}} = \mathrm{Tr}\left(\boldsymbol{x}_j \boldsymbol{x}_j^H\right) = \boldsymbol{x}_j^H \boldsymbol{x}_j \tag{1.5}$$

and each BS will have an independent power constraint

$$P_{j,\mathrm{tx}} \leq P_{j,\max} \tag{1.6}$$

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

$$\boldsymbol{n} = \begin{bmatrix} \boldsymbol{n}_1 \\ \vdots \\ \boldsymbol{n}_N \end{bmatrix} \in \mathbb{C}^{Nr\times 1} \tag{1.7}$$

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 *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 BS in the system so that the information intended for a particular user will be transmitted by several or all the BS. Or, equivalently, each BS transmits a combination of the information of several users

$$\boldsymbol{x}_j = \boldsymbol{W}_{j1}^{(\mathrm{tx})} \boldsymbol{s}_1 + \cdots + \boldsymbol{W}_{jN}^{(\mathrm{tx})} \boldsymbol{s}_N \tag{1.8}$$

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

$$P_{j,\mathrm{tx}} = \sum_{i=1}^{N} \mathrm{Tr}\left(\boldsymbol{W}_{ji}^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_{ji}^{(\mathrm{tx}),H}\right) \tag{1.9}$$

The choice of the precoding matrices and of the receiving filter will determine the transmission strategy used. This Thesis will focus mainly on Block Diagonalization (BD) [1] which is described in Section 1.3.

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

$$\hat{\boldsymbol{s}}_i = \boldsymbol{W}_i^{(\mathrm{rx})} \boldsymbol{y}_i \tag{1.10}$$

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

Combining (1.1)–(1.8) it is possible to rewrite (1.1) as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{W}^{(\mathrm{tx})}\boldsymbol{s} + \boldsymbol{n} \tag{1.11}$$

where

$$\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_N \end{bmatrix} \in \mathbb{C}^{N\ell \times 1} \tag{1.12}$$

and the global precoding matrix is

$$\boldsymbol{W}^{(\mathrm{tx})} = \begin{bmatrix} \boldsymbol{W}_{11}^{(\mathrm{tx})} & \cdots & \boldsymbol{W}_{1N}^{(\mathrm{tx})} \\ \vdots & \ddots & \vdots \\ \boldsymbol{W}_{M1}^{(\mathrm{tx})} & \cdots & \boldsymbol{W}_{MN}^{(\mathrm{tx})} \end{bmatrix} \in \mathbb{C}^{Mt \times N\ell} \tag{1.13}$$

and it can be partitioned as

$$\boldsymbol{W}^{(\mathrm{tx})} = \left[\boldsymbol{W}_1^{(\mathrm{tx})}, \ldots, \boldsymbol{W}_N^{(\mathrm{tx})} \right] \tag{1.14}$$

with

$$\boldsymbol{W}_i^{(\mathrm{tx})} = \begin{bmatrix} \boldsymbol{W}_{11}^{(\mathrm{tx})} \\ \vdots \\ \boldsymbol{W}_{M1}^{(\mathrm{tx})} \end{bmatrix} \in \mathbb{C}^{Mt \times \ell} \tag{1.15}$$

The channel matrix can then be partitioned as

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_N \end{bmatrix} \tag{1.16}$$

where $\boldsymbol{H}_i \in \mathbb{C}^{r \times Mt}$ is the channel matrix from all the BS 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

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{W}^{(\mathrm{tx})} \boldsymbol{s} + \boldsymbol{n}_i \tag{1.17}$$

which can be rewritten as

$$\boldsymbol{y}_i = \boldsymbol{H}_i \boldsymbol{W}_i^{(\mathrm{tx})} \boldsymbol{s}_i + \underbrace{\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{s}_j}_{\text{Interference}} + \boldsymbol{n}_i \tag{1.18}$$

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

$$\boldsymbol{z}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{s}_j + \boldsymbol{n}_i \tag{1.19}$$

and using (1.18), the ergodic (mean) rate for the i-th user is given by [2], [3]

$$R_i = \mathbb{E}\left\{ \log_2 \left| \boldsymbol{I} + \boldsymbol{H}_i \boldsymbol{W}_i^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_i^{(\mathrm{tx}),H} \boldsymbol{H}_i^H \boldsymbol{R}_{\boldsymbol{z}_i}^{-1} \right| \right\} \tag{1.20}$$

where $\boldsymbol{R}_{\boldsymbol{z}_i} \in \mathbb{C}^{r \times r}$ is the covariance matrix of the noise plus the interference term in (1.18)

$$\boldsymbol{R}_{\boldsymbol{z}_i} = \boldsymbol{z}_i \boldsymbol{z}_i^H = \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} + \boldsymbol{n}_i \right) \left(\sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} + \boldsymbol{n}_i \right)^H \tag{1.21}$$

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

1.2 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

$$\boldsymbol{H}_{ij} = \boldsymbol{R}_{\mathrm{rx},ij}\bar{\boldsymbol{H}}_{ij}\boldsymbol{R}_{\mathrm{tx},ij} \tag{1.22}$$

where $\boldsymbol{R}_{\mathrm{tx},ij} \in \mathbb{C}^{t\times t}$ is the spatial correlation at the transmitter, $\bar{\boldsymbol{H}} \in \mathbb{C}^{r\times t}$ is the fading channel, and $\boldsymbol{R}_{\mathrm{rx},ij} \in \mathbb{C}^{r\times r}$ 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 $\boldsymbol{R}_{\mathrm{rx},ij} = \boldsymbol{I}$ and $\boldsymbol{R}_{\mathrm{rx},ij} = \boldsymbol{I}$, as in [4].

The small scale characteristics considered are Rayleigh, so that the entries of $\bar{\boldsymbol{H}}_{ij}$ are independent identically distributed (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 experiments due to the propagation varies according to an exponential power decay with exponent γ, so that the path loss[1] is calculated as

$$\mathrm{pl}_{ij} = \mathrm{pl}_0 \left(\frac{d_{ij}}{d_0}\right)^{-\gamma} \tag{1.24}$$

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.

1.2.1 Signal to Noise Ratio (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 [5], 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 *Rcell*, when the BS transmits at full power, so that the relationship between SNR and the noise power is given by

$$\rho = \frac{P_{\max} Rcell^{-\gamma}}{\sigma_n^2} \tag{1.25}$$

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

$$\mathrm{PL}_{ij}\,(\mathrm{dB}) \triangleq 10\log_{10}\left(\mathrm{pl}_{ij}\right) \tag{1.23}$$

where it is assumed that all the BS have the same maximum transmission power. Using (1.25) it is possible to calculate the noise power for a given SNR and vice versa.

1.3 Block Diagonalization

One possibility to cancel the inter-user interference is to diagonalize the channel matrix. Perfect diagonalization is only possible if $Mt \geq Nr$ [6], and it is achieved using the following precoding matrix

$$\boldsymbol{W}^{(\mathrm{tx})} = \boldsymbol{H}^{\dagger} \tag{1.26}$$

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 [1] is stated that the optimum solution under the constraint that all inter-user interference be zero is obtained with $\boldsymbol{H}\boldsymbol{W}^{(\mathrm{tx})}$ being block diagonal. In [1], 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 (1.18), and this is equivalent to meet the following

$$\boldsymbol{H}_i \boldsymbol{W}_j^{(\mathrm{tx})} = \boldsymbol{0} \quad \forall j \neq i \tag{1.27}$$

Let $\widetilde{\boldsymbol{H}}_i$ be the channel matrix $\boldsymbol{H}$ without $\boldsymbol{H}_i$, *i.e.*

$$\widetilde{\boldsymbol{H}}_i = \begin{bmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_{i-1} \\ \boldsymbol{H}_{i+1} \\ \vdots \\ \boldsymbol{H}_N \end{bmatrix} \in \mathbb{C}^{(N-1)r \times Mt} \tag{1.28}$$

Then the condition (1.27) 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.* $\mathrm{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) > 0$.

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

$$\mathrm{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) = Mt - \mathrm{rank}\left(\widetilde{\boldsymbol{H}}_i\right) \tag{1.29}$$

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

$$\text{rank}\left(\ker\left(\widetilde{\boldsymbol{H}}_i\right)\right) = Mt - \widetilde{L}_i > 0 \tag{1.30}$$

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

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

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

$$\widetilde{\boldsymbol{H}}_i = \widetilde{\boldsymbol{U}}_i \widetilde{\boldsymbol{\Lambda}}_i \left[\widetilde{\boldsymbol{V}}_i^{(1)}, \widetilde{\boldsymbol{V}}_i^{(0)}\right]^H \tag{1.31}$$

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

$$\widetilde{\boldsymbol{H}}_i \widetilde{\boldsymbol{V}}_i^{(0)} = \boldsymbol{0} \tag{1.32}$$

Using these matrices as precoding the result is

$$\widehat{\boldsymbol{H}} = \boldsymbol{H}\left[\widetilde{\boldsymbol{V}}_1^{(0)}, \ldots, \widetilde{\boldsymbol{V}}_N^{(0)}\right] = \begin{bmatrix} \boldsymbol{H}_1 \widetilde{\boldsymbol{V}}_1^{(0)} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{H}_N \widetilde{\boldsymbol{V}}_N^{(0)} \end{bmatrix} \tag{1.33}$$

that, as it can be seen, has a block diagonal structure, which gives the name to the algorithm proposed in [1].

The next problem that BD solves is the maximization of the sum rate of the system, given the block diagonal structure in (1.33). The precoding matrix $\boldsymbol{W}_i^{(\text{tx})}$ will be considered to be

$$\boldsymbol{W}_i^{(\text{tx})} = \widetilde{\boldsymbol{V}}_i^{(0)} \boldsymbol{W}_i' \tag{1.34}$$

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

Introducing (1.34) into (1.20) the ergodic capacity simplifies to

$$R_i^{\text{no interf}} = \mathbb{E}\left\{\log_2\left|\boldsymbol{I} + \frac{1}{\sigma_i^2}\widehat{\boldsymbol{H}}_i \boldsymbol{W}_i' \boldsymbol{R}_{\boldsymbol{s}_i} \boldsymbol{W}_i'^{,H} \widehat{\boldsymbol{H}}_i^H\right|\right\} \tag{1.35}$$

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

$$\widehat{\boldsymbol{H}}_i = \widehat{\boldsymbol{U}}_i \begin{bmatrix} \widehat{\boldsymbol{\Lambda}}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \left[\widehat{\boldsymbol{V}}_i^{(1)}, \widehat{\boldsymbol{V}}_i^{(0)}\right]^H \tag{1.36}$$

where $\widehat{\boldsymbol{\Lambda}}_i = \text{diag}\left\{\hat{\lambda}_{i1}^{1/2}, ..., \hat{\lambda}_{ir}^{1/2}\right\} \in \mathbb{C}^{r\times r}$ 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

$$\boldsymbol{W}_i^{(\text{tx})} = \widetilde{\boldsymbol{V}}_i^{(0)} \widehat{\boldsymbol{V}}_i^{(1)} \tag{1.37}$$

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

$$\boldsymbol{W}_i^{(\text{rx})} = \widehat{\boldsymbol{U}}_i^H \tag{1.38}$$

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

$$R_i^{\text{BD}} = \mathbb{E}\left\{\sum_{k=1}^{\ell} \log_2\left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2}\right)\right\} \tag{1.39}$$

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 1.4.

1.4 Power Allocation

In Section 1.3, 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 (1.39) 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 [7], [8].

- 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 1.4.3 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 BS.

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

$$\begin{aligned} \underset{\left\{\boldsymbol{R}_{\boldsymbol{s}_i}\right\}}{\text{maximize}} \quad & f\left(R_i^{\mathrm{BD}}\left(\boldsymbol{R}_{\boldsymbol{s}_1},\right),\ldots,R_N^{\mathrm{BD}}\left(\boldsymbol{R}_{\boldsymbol{s}_N}\right)\right) \\ \text{subject to} \quad & P_{j,\mathrm{tx}} \leq P_{j,\max}, \quad j=1,\ldots,M \end{aligned} \tag{1.40}$$

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

$$f\left(R_i^{\mathrm{BD}},\ldots,R_N^{\mathrm{BD}}\right)=\sum_{i=1}^{N}\alpha_i R_i^{\mathrm{BD}} \tag{1.41}$$

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

$$\sum_{i=1}^{N}\alpha_i = 1 \tag{1.42}$$

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

$$\overline{\boldsymbol{W}}_j^{(\mathrm{tx})}=\left[\boldsymbol{W}_{j1}^{(\mathrm{tx})},\ldots,\boldsymbol{W}_{jN}^{(\mathrm{tx})}\right]\in\mathbb{C}^{t\times N\ell} \tag{1.43}$$

the precoding matrix of the j-th BS, and

$$\boldsymbol{R}_{\boldsymbol{s}} = \text{blkdiag}\left(\boldsymbol{R}_{\boldsymbol{s}_1}, \dots, \boldsymbol{R}_{\boldsymbol{s}_N}\right) \in \mathbb{C}^{N\ell \times N\ell} \tag{1.44}$$

the matrix containing the power assigned to all the streams of all the users, the power constraint in (1.40) can then be reformulated as

$$\text{Tr}\left(\overline{\boldsymbol{W}}_j^{(\text{tx})} \boldsymbol{R}_{\boldsymbol{s}} \overline{\boldsymbol{W}}_j^{(\text{tx}),H}\right) \leq P_{j,\max} \tag{1.45}$$

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

$$\overline{\boldsymbol{W}}_j^{(\text{tx})} = \left[\bar{\boldsymbol{w}}_{j,11}, \dots, \bar{\boldsymbol{w}}_{j,1\ell}, \dots, \bar{\boldsymbol{w}}_{j,N\ell}\right] \tag{1.46}$$

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

$$\begin{aligned}
\overline{\boldsymbol{W}}_j^{(\text{tx})} \boldsymbol{R}_{\boldsymbol{s}} \overline{\boldsymbol{W}}_j^{(\text{tx}),H} &= \left[\bar{\boldsymbol{w}}_{j,11}, \dots, \bar{\boldsymbol{w}}_{j,N\ell}\right] \begin{bmatrix} p_{11} & & 0 \\ & \ddots & \\ 0 & & p_{N\ell} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{w}}_{j,11}^H \\ \vdots \\ \bar{\boldsymbol{w}}_{j,N\ell}^H \end{bmatrix} = \\
&= p_{11}\bar{\boldsymbol{w}}_{j,11}\bar{\boldsymbol{w}}_{j,11}^H + \cdots + p_{N\ell}\bar{\boldsymbol{w}}_{j,N\ell}\bar{\boldsymbol{w}}_{j,N\ell}^H \\
&= \sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\bar{\boldsymbol{w}}_{j,ik}\bar{\boldsymbol{w}}_{j,ik}^H
\end{aligned} \tag{1.47}$$

and the trace is

$$\text{Tr}\left(\overline{\boldsymbol{W}}_j^{(\text{tx})} \boldsymbol{R}_{\boldsymbol{s}} \overline{\boldsymbol{W}}_j^{(\text{tx}),H}\right) = \sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 \tag{1.48}$$

where

$$\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 = \text{Tr}\left(\bar{\boldsymbol{w}}_{j,ik}\bar{\boldsymbol{w}}_{j,ik}^H\right) = \bar{\boldsymbol{w}}_{j,ik}^H\bar{\boldsymbol{w}}_{j,ik}. \tag{1.49}$$

The sum-rate maximization problem can then be formulated in *standard form* [9] as

$$\begin{aligned}
\underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right) & \\
\text{subject to} \quad & \sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max} \leq 0, & j = 1,\dots,M \\
& -p_{ik} \leq 0, & \begin{matrix} i = 1,\dots,N \\ k = 1,\dots,\ell \end{matrix}
\end{aligned} \tag{1.50}$$

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

1.4.1 Optimal Power Allocation

In (1.50), the function $\log_2(\cdot)$ is convex on p_{ik} and the sum of convex functions is also convex, so the objective function in (1.50) is convex. The constraints are affine and therefore convex too. The optimization problem in (1.50) is a convex optimization problem that can be solved using a myriad of numerical technics [9]. Nonetheless, it would be interesting to analyze a bit further the problem in order to get some insight about it.

The problem in (1.50) satisfies *Slater's condition* [9] 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 [9] 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 (1.50) are

$$\begin{aligned}
\sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}^{*}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max} &\leq 0, \quad j=1,\dots,M \\
-p_{ik}^{*} &\leq 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
-\nu_j^{*} &\geq 0, \quad j=1,\dots,M \\
-\mu_{ik}^{*} &\geq 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
\nu_j^{*}\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell} p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 - P_{j,\max}\right) &= 0, \quad j=1,\dots,M \\
-\mu_{ik}^{*}p_{ik}^{*} &= 0, \quad \begin{matrix} i=1,\dots,N \\ k=1,\dots,\ell \end{matrix} \\
\boldsymbol{\nabla}_{\boldsymbol{p}}\mathcal{L}\left(\boldsymbol{p}^{*},\boldsymbol{\nu}^{*},\boldsymbol{\mu}^{*}\right) &= \mathbf{0}
\end{aligned} \tag{1.51}$$

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 (1.50), and it is defined as

$$\begin{aligned}\mathcal{L}\left(\boldsymbol{p},\boldsymbol{\nu},\boldsymbol{\mu}\right) = &-\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right)+\\ &\sum_{j=1}^{M}\nu_j\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2-P_{j,\max}\right)-\\ &\sum_{i=1}^{N}\sum_{k=1}^{\ell}\mu_{ik}p_{ik}\end{aligned} \tag{1.52}$$

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

$$\boldsymbol{\nabla}_{\boldsymbol{x}}f(\boldsymbol{x}) = \begin{bmatrix}\frac{\partial}{\partial x_1}f(\boldsymbol{x})\\ \vdots \\ \frac{\partial}{\partial x_n}f(\boldsymbol{x})\end{bmatrix} \tag{1.53}$$

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

$$\frac{\partial}{\partial p_{ik}}\left\{-\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right)\right\} = \frac{-\hat{\lambda}_{ik}}{\ln\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}\right)} \tag{1.54}$$

And the same for the inequality constraints

$$\begin{aligned}&\frac{\partial}{\partial p_{ik}}\left\{\sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2-P_{j,\max}\right\} = \left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2\\ &\frac{\partial}{\partial p_{ik}}\left\{p_{ik}\right\} = 1\end{aligned} \tag{1.55}$$

So that the condition of the gradient of the Lagrangian vanishing, in (1.50) can be writen as

$$\frac{-\hat{\lambda}_{ik}}{\ln\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}^*\right)}+\sum_{j=1}^{M}\nu_j^*\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2-\mu_{ik}^* = 0, \qquad \begin{matrix}i=1,\dots,N\\ k=1,\dots,\ell\end{matrix} \tag{1.56}$$

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

$$\frac{-\hat{\lambda}_{ik}}{\ln\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}^*\right)}+\sum_{j=1}^{M}\nu_j^*\left\|\bar{\boldsymbol{w}}_{j,ik}\right\|_2^2 \geq 0, \qquad \begin{matrix}i=1,\dots,N\\ k=1,\dots,\ell\end{matrix} \tag{1.57}$$

Calling

$$L_{ik} = \sum_{j=1}^{M} \nu_j^* \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \tag{1.58}$$

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

$$p_{ik}^* \leq \frac{1}{\ln(2)\, L_{ik}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}}, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \tag{1.59}$$

The result in (1.59) 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 BS 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.

1.4.2 Modified Water-Filling

[10] proposes a simplification to the original problem, in order to make it more tractable. The coupling of the power constraints in (1.50) makes it impossible to get a simple solution for the optimal power allocation problem. [10] 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 BS consider a single power constraint given by the most restrictive BS in the original problem. Define

$$\Omega_{ik}^{\mathrm{BS}} = \max_{j=1,\ldots,M} \left\| \bar{\boldsymbol{w}}_{j,ik} \right\|_2^2 \tag{1.60}$$

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

$$\begin{aligned} \underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N} \sum_{k=1}^{\ell} \log_2 \left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2} \right) \\ \text{subject to} \quad & \sum_{i=1}^{N} \sum_{k=1}^{\ell} p_{ik} \Omega_{ik}^{\mathrm{BS}} - P_{\mathrm{BS,max}} \leq 0 \\ & -p_{ik} \leq 0, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \end{aligned} \tag{1.61}$$

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

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

$$\begin{aligned}\mathcal{L}\left(\boldsymbol{p},\nu,\boldsymbol{\mu}\right) = &-\sum_{i=1}^{N}\sum_{k=1}^{\ell}\log_2\left(1+\frac{\hat{\lambda}_{ik}p_{ik}}{\sigma_i^2}\right)+\\ &\nu\left(\sum_{i=1}^{N}\sum_{k=1}^{\ell}p_{ik}\Omega_{ik}^{\mathrm{BS}}-P_{\mathrm{BS,max}}\right)-\\ &\sum_{i=1}^{N}\sum_{k=1}^{\ell}\mu_{ik}p_{ik}\end{aligned} \tag{1.62}$$

And its gradient is given by

$$\boldsymbol{\nabla}_{\boldsymbol{p}}\mathcal{L}\left(\boldsymbol{p}^*,\nu^*,\boldsymbol{\mu}^*\right) = \frac{-\hat{\lambda}_{ik}}{\ln\left(2\right)\left(\sigma_i^2+\hat{\lambda}_{ik}p_{ik}^*\right)}+\nu^*\Omega_{ik}^{\mathrm{BS}}-\mu_{ik}^* = 0, \qquad \begin{matrix} i=1,\ldots,N\\ k=1,\ldots,\ell\end{matrix} \tag{1.63}$$

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

$$p_{ik}^* \leq \frac{1}{\ln\left(2\right)\nu^*\Omega_{ik}^{\mathrm{BS}}}-\frac{\sigma_i^2}{\hat{\lambda}_{ik}}, \qquad \begin{matrix} i=1,\ldots,N\\ k=1,\ldots,\ell\end{matrix} \tag{1.64}$$

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

$$p_{ik}^* = \left[\frac{1}{\ln\left(2\right)\nu^*\Omega_{ik}^{\mathrm{BS}}}-\frac{\sigma_i^2}{\hat{\lambda}_{ik}}\right]^+, \qquad \begin{matrix} i=1,\ldots,N\\ k=1,\ldots,\ell\end{matrix} \tag{1.65}$$

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 [11].

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 [10] 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 [10] 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

$$p_{ik}^{*} = \left[K_{\mathrm{WF}} - \frac{\sigma_i^2}{\hat{\lambda}_{ik}} \right]^{+}, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \tag{1.66}$$

where $K_{\mathrm{WF}} = \frac{1}{\ln(2)\nu^{*}\Omega_{\mathrm{WF}}}$, 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.

1.4.3 Scaled Water-Filling

In [5] the same power allocation problem as in (1.50) is dealt with by considering a TPC, so that the optimization problem becomes

$$\begin{aligned} \underset{p_{ik}}{\text{minimize}} \quad & -\sum_{i=1}^{N}\sum_{k=1}^{\ell} \log_2\left(1 + \frac{\hat{\lambda}_{ik} p_{ik}}{\sigma_i^2}\right) \\ \text{subject to} \quad & \mathrm{Tr}\left(\boldsymbol{W}^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}} \boldsymbol{W}^{(\mathrm{tx}),H}\right) - M P_{\max} \leq 0 \\ & -p_{ik} \leq 0, \qquad \begin{matrix} i = 1, \ldots, N \\ k = 1, \ldots, \ell \end{matrix} \end{aligned} \tag{1.67}$$

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

Under this TPC, the solution is readily derived by water-filling [11], 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

$$\boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{SWF}} = \beta \boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}} \tag{1.68}$$

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

$$\beta = \frac{P_{\max}}{\max\limits_{j=1,\ldots,M} \mathrm{Tr}\left(\boldsymbol{W}_j^{(\mathrm{tx})} \boldsymbol{R}_{\boldsymbol{s}}^{\mathrm{TPC}} \boldsymbol{W}_j^{(\mathrm{tx}),H}\right)} \tag{1.69}$$

The results in [5] show, as well, that this simplified approach can deliver near-optimum performance.

1.4.4 Uniform Power allocation

The simplest, both conceptually and computationally, alternative that can be considered to solve the power allocation in (1.50) 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

$$p_{ik} = p_s, \qquad \begin{matrix} i = 1, \dots, N \\ k = 1, \dots, \ell \end{matrix} \tag{1.70}$$

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

Recall from (1.45) the power transmitted by the j-th BS, where now the matrix $\boldsymbol{R_s}$ is given by

$$\boldsymbol{R_s} = p_s \boldsymbol{I} \tag{1.71}$$

and (1.45) becomes

$$p_s \operatorname{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right) \leq P_{j,\max}, \quad j = 1, \dots, M \tag{1.72}$$

The new power allocation problem can be formulated as

$$\begin{aligned} &\underset{p}{\text{maximize}} \quad p \\ &\text{subject to} \quad p \operatorname{Tr}\left(\overline{\boldsymbol{W}}_j^{(\mathrm{tx})} \overline{\boldsymbol{W}}_j^{(\mathrm{tx}),H}\right) \leq P_{j,\max}, \qquad j = 1, \dots, M \end{aligned} \tag{1.73}$$

which is a *linear programming* optimization problem, and it can be solved efficiently using classical methods, *e.g.* bisection method [12].

Chapter 2

Mean Achievable Rates in Clustered Coordinated Base Station Transmission with Block Diagonalization[1]

2.1 Introduction

In \refc{ch:state_art} 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 chose 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.

2.2 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

[1]The work shown in this chapter has been published in [13].

Figure 2.1: 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 BS.

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 [14], 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 BS 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 BS in the cluster are considered cluster members, that is, no scheduling or adaptive selection of active BS 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 2.1 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.

2.3 Interference Model

In Figure 2.1 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 $d_i \triangleq d_{ii}$ from its serving BS, 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 $d_j \triangleq d_{jj}$ from its serving BS, 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 2.1, 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

$$f_{d_i}(d_i) = \frac{2d_i}{R_{\text{cell}}^2} \tag{2.1}$$

The interference power received at the user i is equal to

$$I_i(d_i) = \sum_{m=1}^{M_{\text{interf}}} P_{\max} \bar{d}_{im}^{-\gamma} \tag{2.2}$$

where M_{interf} is the number of interfering BS, *i.e.* the total number of cells in the system minus the M cells that form the cluster, and $\bar{d}_{im}$ is the distance from the m-th BS outside the cluster to the i-th user in the cluster. All the interfering BS 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

$$D_i \in \{D_{\text{tier 1}}, D_{\text{tier 2}}\} \tag{2.3}$$

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

With all this, (2.2) can be written as

$$I_i(d_i) = P_{\max} M_{\text{eq},i} (D_i - d_i)^{-\gamma} \tag{2.4}$$

This approach is similar to that followed in [15], where a fluid model network is used. This model assumes that there is a continuum of BS interfering, but this will not be considered in the current work for the sake of simplicity.

Figure 2.2: Cluster with $M = 4$ cells and neighbor interfering cells.

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

In order to determine $M_{\text{eq},i}$ 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 BS for each of the cells in the cluster, and this number for the i-th cell is denoted as $M_{\text{int},i} \leq M_{\text{interf}}$.

This is made clear in Figure 2.2 where a cluster with $M = 4$ cells is surrounded by $M_{\text{interf}} = 12$ cells. In this cluster, cells 1 and 3 experience an interference coming from $M_{\text{int},1} = M_{\text{int},3} = 4$ neighboring cells, while cells 2 and 4 receive the interference from $M_{\text{int},2} = M_{\text{int},4} = 3$ cells.

In order to approximate the value of equivalent interfering BS for a general network setup, fist consider a simple scenario Figure 2.3 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 $M_{\text{int},1} = 6$ belonging to the first tier.

An assumption that can be made is that half of the $M_{\text{int},1}$, *i.e.* three, BS 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

$$I_1 = P_{\max}\left[\frac{3}{(D_1 + d_1)^{\gamma}} + \frac{3}{(D_1 + d_1)^{\gamma}}\right] \tag{2.5}$$

which is equal to (2.4) if the equivalent number of interfering BS is defined as

$$M_{\text{eq},1} = 3\left[1 + \left(\frac{D_1 - d_1}{D_1 + d_1}\right)^{\gamma}\right] \tag{2.6}$$

In general, if a user i is considered to have $M_{\text{int},i}$ interfering BS in the first tier at a distance D_i, defined as in (2.3), then the equivalent number of interfering BS is given by

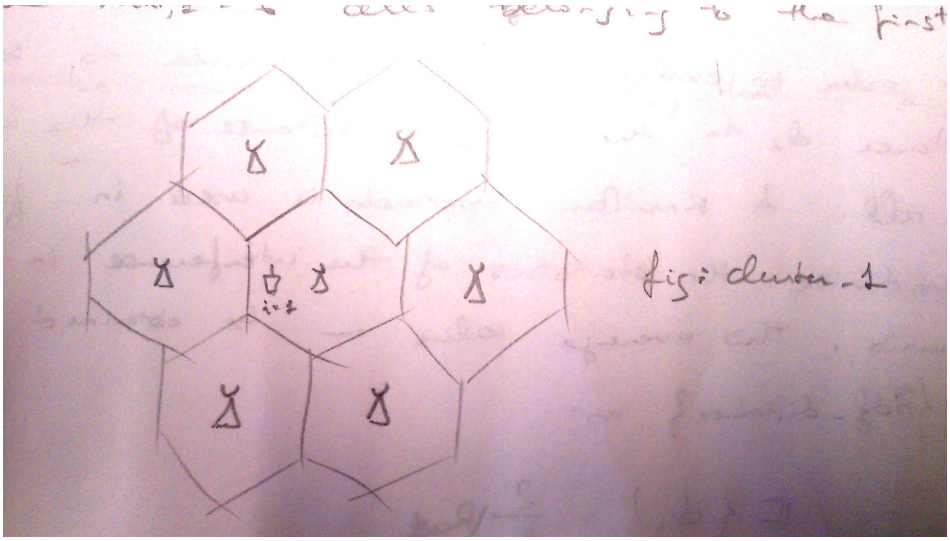


Figure 2.3: Simple scenario with a single cluster of $M = 1$ cell, and the six interfering cells surrounding it.

$$M_{\mathrm{eq},i} \approx \frac{M_{\mathrm{int},i}}{2}\left[1+\left(\frac{D_i-d_i}{D_i+d_i}\right)^{\gamma}\right] \tag{2.7}$$

In order to evaluate (2.7), 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 [16] to characterize the statistics of the interference in a multi-cell scenario. This average distance can be obtained from the uniform spatial distribution in (2.1) as

$$\mathbb{E}\{d_i\} = \frac{2}{3}R_{\mathrm{cell}} \tag{2.8}$$

and then

$$M_{\mathrm{eq},i} \approx \frac{M_{\mathrm{int},i}}{2}\left[1+\left(\frac{D_i-\frac{2}{3}R_{\mathrm{cell}}}{D_i+\frac{2}{3}R_{\mathrm{cell}}}\right)^{\gamma}\right] \tag{2.9}$$

Section 2.5 deals with the comparison between simulations and the analytical results, showing that the approximation in (2.9) is rather good.

2.4 Analysis of the Rate

2.5 Numerical Results

Bibliography

[1] 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, pp. 461–471, February 2004.

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

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

[4] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *Wireless Communications, IEEE*, vol. 13, pp. 56–61, Aug 2006.

[5] 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, pp. 1910–1921, April 2009.

[6] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, pp. 1691–1706, July 2003.

[7] 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, pp. 69–73, Jan 2007.

[8] 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, pp. 1076–1086, March 2012.

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

[10] 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.

[11] J. M. Cioffi, *EE379C - Advanced Digital Communication - Class Notes - Chapter 4*.

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

[13] 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, pp. 3483–3493, August 2013.

[14] 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, pp. 6086–6101, Dec 2011.

[15] 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.

[16] 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.