
# ANTENNA SELECTION IN MASSIVE MIMO BASED ON MATCHING PURSUIT

Marcele Oliveira Kuhfuss de Mendonça

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientadores: Paulo Sergio Ramirez Diniz
Tadeu Nagashima Ferreira

Rio de Janeiro
Agosto de 2018

ANTENNA SELECTION IN MASSIVE MIMO BASED ON MATCHING PURSUIT

Marcele Oliveira Kuhfuss de Mendonça

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Examinada por:

Prof. Paulo Sergio Ramirez Diniz, Ph.D.

Prof. Tadeu Nagashima Ferreira, D.Sc.

Prof. Wallace Alves Martins, D.Sc.

Prof. Lisandro Lovisolo, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
AGOSTO DE 2018

Mendonça, Marcele Oliveira Kuhfuss de

Antenna Selection in Massive MIMO based on Matching Pursuit/Marcele Oliveira Kuhfuss de Mendonça. – Rio de Janeiro: UFRJ/COPPE, 2018.

XVII, 97 p.: il.; 29,7cm.

Orientadores: Paulo Sergio Ramirez Diniz

Tadeu Nagashima Ferreira

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2018.

Referências Bibliográficas: p. 88 – 97.

1. massive MIMO. 2. antenna selection. 3. matching pursuit. I. Diniz, Paulo Sergio Ramirez *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*To my family*

# Agradecimentos

Agradeço especialmente às três pessoas mais importantes da minha vida: meus pais Maria Cláudia e Marcelo e meu irmão Marcel. Obrigada por me motivarem a seguir em frente mesmo nas adversidades. Obrigada por serem muito mais que uma família. Vocês são tudo para mim. Mãe, obrigada, principalmente por seus sábios conselhos, e por me mostrar que sonhar ainda vale a pena. Não sei o que seria de mim sem você. Pai, obrigada, principalmente por me motivar sempre, e por estar presente com o seu jeitinho alegre e admirável. Marcel, obrigada por seu companheirismo e positivismo, e por ainda acreditar nos nossos sonhos. Agradeço ao meu irmão Matheus. Sei que você torce por mim de onde você está. Obrigada pelas boas energias que manda para mim e para nossa família. Agradeço também ao meu filhotinho peludo Simba, por ficar ao meu lado nas madrugadas longas de textos e códigos, inclusive nessa em que estou escrevendo estas palavras. Agradeço ao meu tio Luiz Carlos e meus avós Antônia e Raimundo, obrigada por sempre poder contar com vocês e por estarem sempre torcendo por mim.

Agradeço ao meus orientadores Diniz e Tadeu. Vocês são as duas pessoas que não fazem parte da minha família de sangue que eu mais admiro, respeito e tenho um profundo carinho. Diniz, obrigada por sempre acreditar no meu potencial, mesmo quando nem eu acreditava. Obrigada por se preocupar não só com o desenvolvimento deste trabalho, mas também com o meu bem estar emocional. Obrigada por suas palavras amigas e pelos seus conselhos. Obrigada por sempre estar disposto a me ouvir e ajudar. Tadeu, obrigada por ser meu orientador mais uma vez. Obrigada por seus ensinamentos, pelo seu acompanhamento detalhado e impecável deste trabalho. Obrigada por me motivar a ser uma profissional melhor. Obrigada por ser um grande amigo, por me ouvir e me ajudar sempre que você pôde ao longo desses anos que te conheço.

Agradeço aos colegas do SMT por toda ajuda direta ou indireta, companheirismo nas matérias, e cafés animados.

Agradeço aos professores Wallace e Lisandro por aceitarem fazer parte da minha banca de mestrado.

Agradeço também ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro durante o mestrado.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

# SELEÇÃO DE ANTENAS EM SISTEMA MASSIVE MIMO BASEADA EM MATCHING PURSUIT

Marcele Oliveira Kuhfuss de Mendonça

Agosto/2018

Orientadores: Paulo Sergio Ramirez Diniz
Tadeu Nagashima Ferreira

Programa: Engenharia Elétrica

Com a proliferação de serviços *wireless*, a demanda por espectro disponível também cresce. Logo, a eficiência espectral é um assunto de grande interesse na comunidade científica, que procura por meios para fornecer qualidade de serviço ao crescente número de usuários. *massive MIMO* é uma técnica repleta de atrativos a ser empregada na futura geração *wireless*, já que aproveita o espectro existente eficientemente. Este trabalho propõe duas estratégias de seleção de antenas para serem empregadas no *downlink* de um sistema *massive MIMO*, visando a redução da potência de transmissão. Os algoritmos propostos podem também ser usados para selecionar um subconjunto de sensores ativos em uma rede centralizada de sensores. A estratégia proposta para seleção de antenas é inspirada na técnica *matching pursuit*. Os resultados apresentados indicam que uma seleção eficiente pode ser obtida com baixa complexidade computacional.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

# ANTENNA SELECTION IN MASSIVE MIMO BASED ON MATCHING PURSUIT

Marcele Oliveira Kuhfuss de Mendonça

August/2018

Advisors: Paulo Sergio Ramirez Diniz
Tadeu Nagashima Ferreira

Department: Electrical Engineering

As wireless services proliferate, the demand for available spectrum also grows. As a result, the spectral efficiency is still an issue addressed by many researchers looking for solutions to provide quality of service to a growing number of users. massive MIMO is an attractive technology for the next wireless systems since it can alleviate the expected spectral shortage. This work proposes two antenna selection strategies to be applied in the downlink of a massive MIMO system, aiming at reducing the transmission power. The proposed algorithms can also be employed to select a subset of active sensors in centralized sensor networks. The proposed strategy to select the antennas is inspired by the matching pursuit technique. The presented results show that an efficient selection can be obtained with reduced computational complexity.

# Contents

# List of Figures

# List of Tables

# List of Symbols

| | |
|---|---|
| $\mathbf{a}$ | array response or steering vector |
| $\alpha$ | path-loss exponent |
| $\mathbf{b}$ | target vector in MP |
| $\beta_k$ | large-scale fading between terminal $k$ and the BS |
| $\mathbf{c}$ | codeword vector in MP |
| $\chi$ | Schur complement |
| $\mathbf{D}$ | dictionary matrix in MP |
| $\boldsymbol{\eta}$ | vector comprising the power allocated for each user |
| $\Gamma$ | parameter that determines the median channel gain at a reference distance of 1 km |
| $\mathbf{G}$ | channel matrix between BS and terminal(s) |
| $\mathbf{G}_S$ | $S$-selected channel matrix |
| $\kappa$ | positive constant which controls the quality of the barrier approximation |
| $\mathcal{E}$ | confidence ellipsoid of the estimation error vector |
| $\mathcal{H}$ | Hilbert space |
| $\mathcal{I}$ | set containing auxiliary antenna's indices |
| $\mathcal{I}$ | set containing selected antenna's indices |
| $\mathbf{P}$ | precoding matrix |
| $\mathbf{p}$ | precoding vector |
| $\mathbf{P}_S$ | $S$-selected precoding matrix |
| $\mathbf{\Phi}$ | pilot matrix |

$\boldsymbol{\phi}_i$ pilot sequence assigned to terminal $i$

$\mathbf{q}$ intended transmitted symbols

$\mathbf{r}$ residue vector in MP

$\mathbf{R}_k$ spatial correlation matrix

$\rho_{\mathrm{dl}}$ downlink SNR

$\rho_{\mathrm{ul}}$ uplink SNR

$\sigma_w^2$ noise variance

$\tau_{\mathrm{c}}$ length of coherence interval in samples

$\tau_{\mathrm{dl}}$ samples reserved for downlink data transmission in each coherence block

$\tau_{\mathrm{p}}$ samples reserved for uplink pilot signaling in each coherence block

$\tau_{\mathrm{ul}}$ samples reserved for uplink data transmission in each coherence block

$\theta$ angle between the location of the terminal/obstacle and the ground

$\mathbf{V}$ decoding matrix

$\mathbf{v}$ decoding vector

$\mathbf{w}$ AWGN noise

$\mathbf{x}$ transmitted signal

$\mathbf{X}_{\mathrm{p}}$ pilot signal transmitted by all terminals

$\mathbf{y}$ received signal

$\mathbf{Z}$ antenna selection matrix

$\mathbf{z}$ selection vector

$B_c$ coherence bandwidth

$C_{\mathrm{dl}}$ downlink capacity

$C_{\mathrm{ul}}$ uplink capacity

$d_H$ antenna spacing, measured in the number of wavelengths

$d_k$ distance between the BS and the terminal

| | |
|---|---|
| $f_k$ | realization of a random variable which models the shadow fading effect |
| $g_{mk}$ | channel gain between the $k$th terminal and the $m$th base station antenna |
| $h_{mk}$ | small-scale fading between terminal $k$ and the $m$th BS antenna |
| $K$ | number of terminals, except in SU MIMO where $K$ is the number of antennas in the terminal |
| $M$ | number of Base Station antennas |
| $N$ | number of multipaths in NLOS propagation environment |
| $T_c$ | coherence time |
| $\mathbf{g}_{Sk}$ | channel gain between terminal $k$ and $S$ selected BS antennas |
| $\mathbf{g}_m^{\mathrm{T}}$ | channel gain between BS antenna $m$ and the $K$ terminals |
| $\mathbf{g}_k$ | channel gain between terminal $k$ and $M$ BS antennas |

# Chapter 1

# Introduction

Nowadays, it is hard to imagine our lives without cellphones. Interestingly, these devices have become essential in such a short period of time. About 40 years ago, when the first commercial mobile telephones were launched, they were quite expensive and restricted to a small group of people. Now, everyone has a cellphone or at least has access to one. To achieve the cellular systems we have today, the ladder had several steps; these steps are often referred to as cellular generations.

## 1.1 The Road to 5G

The first mobile phone systems appeared soon after the Second World War, but they were not easily portable as one can see in Figure 1.1a. This pre-cellular mobile telephony technology is known as generation zero (0G). The number of simultaneous calls was limited as each call required a separate frequency channel.

The first generation (1G) systems, based on analog technology with FM modulation, appeared in the 1980s and supported speech services only. Although they were still quite big, as shown in Figure 1.1b, they were the first handheld telephones. The US-developed advanced mobile phone service (AMPS) is a typical example of 1G system based on the frequency division multiple access (FDMA) technology which allowed multiple users in a cell. Nevertheless, other countries were also developing their own systems and standards, which were incompatible with each other. In the face of the incompatibility among various systems, changes in the existing telecommunication regulatory framework started to be studied.

In the early 1990s, a transition from analog to digital systems was witnessed, bringing together the second generation (2G), which steadily evolved to support digital voice and also data services such as emails and short messages (SMS). Unlike 1G systems, in which one call required one frequency channel, in 2G systems one frequency channel was divided between several users thanks to time-division or code-division techniques. Therefore, the 2G systems enabled the accommodation of more


(a) A radio telephone from 0G [2].

(b) A "brick" cellphone from 1G [3].

Figure 1.1: Old generation mobile telephones.

subscribers in the radio spectrum. Examples of 2G systems include global system mobile (GSM), IS-136 which is hybrid time division multiple access (TDMA) and frequency division multiple access (FDMA) and IS-95 which is code division multiple access (CDMA). The most famous of them is undoubtedly the GSM, whose open standard easily allowed interoperability.

Given the use of the 2G mobile system, the use of the Internet as a multimedia service provision has quickly increased. As expected, Internet-based services were incorporated in the mobile devices from the third generation (3G) system. The 3G systems were defined by the International Telecommunication Union (ITU) to support high speed data ranges from 144 kbps to 2 Mbps. Examples of successful 3G standards include the Universal Mobile Telecommunications System (UMTS) and CDMA 2000, which are handled by the Third-generation Partnership Project (3GPP) and 3GPP2, respectively. The UMTS or Wideband Code Division multiple access (WCDMA) is the evolved version of GSM, whereas the CDMA 2000 is the evolution of IS-95. The 3G has also introduced the term "mobile broadband" due to its rate and capability of delivering Internet browsing into the cellphones.

More support to broadband data is provided by the fourth generation (4G) or Long Term Evolution (LTE). Originally, LTE was an enhanced version of the 3G technologies named by 3GPP in 2004 [4]. Although the LTE was not intended to be used as synonym for 4G, its use is widespread both in the engineering community and the general public. In LTE, the air interface technology WCDMA is replaced by the orthogonal frequency division multiple access (OFDMA). Another attempt of improvement used in LTE is the multiple input multiple output (MIMO) technique with arrays containing about eight antennas [5].

As the progress never stops, recent research effort is being conducted in the direction of the fifth generation (5G) [6–9]. Basically, this future generation will be concerned with three aspects: higher data-rates, lower latency, and the ability to connect not only people, but also things [10, 11].

## 1.2 The Big Picture of 5G and Motivation

We are moving towards a future where everything is connected. Over the last years, the cellphones evolved to smartphones – powerful devices containing numerous applications and embedded sensors to provide an experience beyond the initial speech service. As a result, the volume of data is dramatically increasing, as quantified by the annual visual network index (VNI) reports released by Cisco in Figure 1.2a. Not only the smartphones will play a significant role in the new generation wireless communications, but the Machine-to-Machine (M2M) communications [12, 13] are also gaining ground as evidenced in Figure 1.2b. M2M communications include smart power-grid, eHealth, intelligent transportation systems and surveillance, to name a few [13].

The most recent VNI report [1] and forecast indicate that a huge improvement should be achieved in data communication to meet the demands that networks will face by 2021. Ultra-densification, mmWave (milimeter wave), and massive MIMO are being considered "the big three" technologies for 5G [14].

A possible solution to deal with the demand for higher data rate increase is to make the cells smaller, hence densifying the area with more BSs (Base Stations). The benefits include the reuse of spectrum across a geographic area and the decrease in the number of users competing for system resources. Moreover, reducing the size of the cell leads to increased spectral efficiency and reduced transmit power.

Since the spectrum range between 300 MHz and 3 GHz are becoming nearly fully occupied, we need to use higher frequencies to access wider bandwidth. Indeed, between 30 GHz and 300 GHz, where wavelengths are 1–10 mm, the spectrum is less crowded and much wider bandwidths are available. On the other hand, millimeter wave signals are highly attenuated by commonly used solid materials such as concrete or brick walls [15, 16]. A possible solution to this problem involves the placement of mmWave *femtocells* inside buildings for indoor coverage.

Another approach is the use of a large array of antenna elements to provide diversity and compensate for path loss. The array is supposed to be equipped with hundreds of antenna elements. Indeed, massive MIMO is one of the most promising technologies for the future-generation wireless systems [17, 18]. In such a strategy, the base station is equipped with a large number of antennas $M$ and serves a set of $K$ terminals or user equipment. One of the characteristics of massive MIMO is

that $M \gg K$ which brings about the favorable action of the law of large numbers [19]. This effect is known as channel hardening, in which all small-scale effects vanish as the number of antennas increases. However, to maintain a certain level of performance, all the $M$ antennas in BS do not need to be active at the same time. Moreover, simple linear processing is allowed due to the quasi-orthogonal nature of the channels between each BS and the set of active users. Nonetheless, we must overcome several issues so that massive MIMO can become a reality.

One of the challenges for the realization of the massive MIMO is the BS cost. In general, the massive MIMO architecture requires radio frequency (RF) elements for each antenna. Thus, increasing the number of antennas leads to an increase in the number of these elements. Motivated by this issue, we aim at selecting the most effective BS antennas in order achieve a certain performance level at the receivers, leading to power savings as well. Therefore we can alleviate the BS cost and also benefit from significant diversity gain.

Solving the problem of selecting $S$ out of $M$ available antennas by verifying all possible choices is quite a challenge. This problem can be solved via convex optimization as shown in [20, 21], but this solution leads to high computational cost. Therefore, we propose an efficient way to solve the antenna selection problem by using a greedy algorithm called matching pursuit [22, 23], in which the computational cost is substantially reduced.

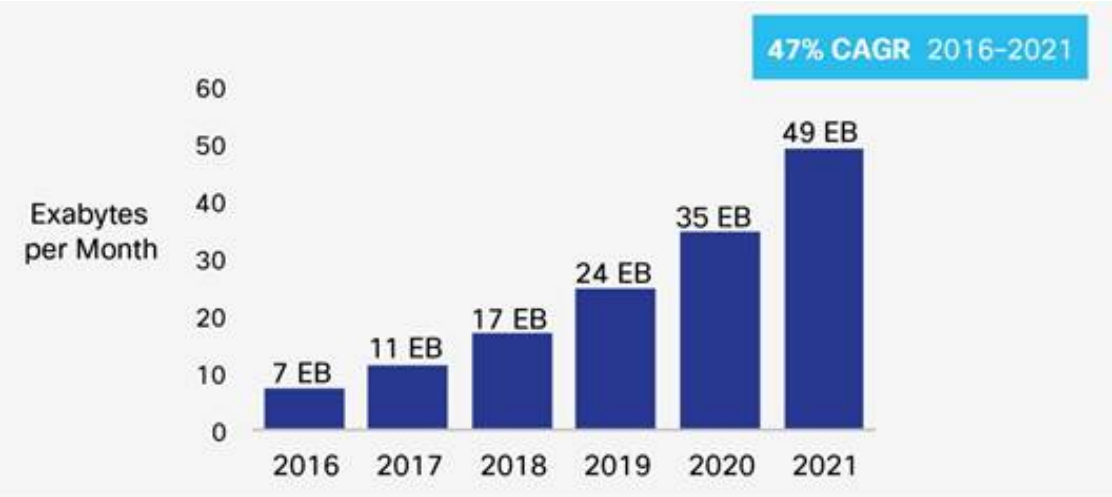
## 1.3 Organization

The structure of the dissertation is as follows. In chapter 2, cellular networks are briefly described as well as the main concepts regarding wireless systems. This chapter also states the possible ways to improve the area throughput of cellular networks, which culminates in the use of multiple antennas at the BS. The MIMO technology is also introduced in chapter 2.

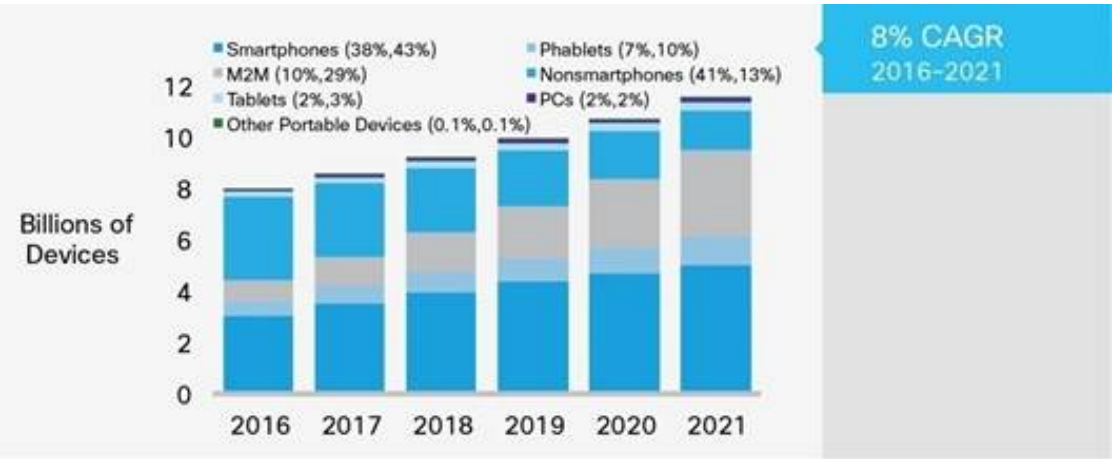
The massive MIMO concept is detailed in chapter 3. This chapter embraces the system model employed and also the main challenges that we need to overcome in massive MIMO. In addition, the concept of channel hardening and favorable propagation is also explored in chapter 3.

The antenna selection problem is described in chapter 4. This chapter states the problem of selecting the best subset of BS antennas for massive MIMO systems. Moreover, chapter 4 describes the existing antenna selection algorithms which are based on convex optimization.

Chapter 5 discusses the matching pursuit technique, a greedy approach used in the sparse recovery context. This chapter contains the proposed antenna selection algorithms which are based on matching pursuit strategy, the main contribution




(a) Global Mobile Data Traffic, 2016 to 2021.




(b) Global Mobile Devices and Connections Growth, 2016 to 2021.

Figure 1.2: Source: Cisco VNI Mobile, 2017. [1]

of this work. We propose both channel-level and symbol-level antenna selection algorithms.

In chapter 6, the proposed antenna selection algorithms and their counterparts are evaluated via bit error rate in a massive MIMO simulated system. The simulations are conducted in different scenarios regarding channel model, modulation scheme and channel estimation. In addition, the computational complexity of the proposed algorithms is quantified by counting the number of flops required to perform the antenna selection. The antenna selection is performed by computing an auxiliary vector that indicates the active antennas indices. The results show that our proposed algorithms achieve a performance level almost as high as the ones yielded by their counterparts based on convex optimization. However, the proposed algorithms have the advantage of requiring a reduced number of flops.

Furthermore, a relation between SU MIMO system and Sensor Networks is addressed in chapter 8. This chapter also includes a simple proposed antenna selection

algorithm based on low complexity approximation.

Finally, some concluding remarks are included in chapter 7 along with possible future work.

## 1.4 Notation

Vectors and matrices are represented by characters in bold type in which lower-case letters are used for vectors and upper-case letters for matrices, whereas non-bold letters are scalar variables. We consider column vectors and $a_m$ represents the $m$th component of vector $\mathbf{a}$. In a similar way, the entries of a matrix $\mathbf{A}$ are of the form $a_{mk}$ in which $m$ represents the row and $k$ the column of $\mathbf{A}$. In order to identify a column of a matrix, we represent it as a column vector $\mathbf{a}_k$ where $k$ is the column index. Similarly, $\mathbf{a}_m^{\mathrm{T}}$ is used to represent the $m$th row of matrix $\mathbf{A}$. For example, an $M \times K$ matrix $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MK} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_K \end{bmatrix}. \tag{1.1}$$

When the elements of a vector are random variables, we represent the vector by a character in bold italic type, i.e., $\boldsymbol{a}$. The statement $\boldsymbol{a} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ means that the random vector $\boldsymbol{a}$ is distributed as a real Gaussian random variable with zero mean and covariance matrix $\mathbf{R}$. The distribution in the statement $\boldsymbol{a} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ is known as circularly simmetric complex Gaussian which means that $e^{j\phi}\boldsymbol{a} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$ for any given $\phi$. E[·] and Var[·] denote the expected value and variance of a random variable, respectively.

In general, we use subscripts in vectors and matrices just to represent the variable's name. However, subscripts in parentheses refer to the size of a square matrix. For example, $\mathbf{I}_{(K)}$ is the identity matrix with size $K \times K$.

The real, complex and natural sets are represented by the following symbols $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{N}$. For example, we can establish that matrix $\mathbf{A} \in \mathbb{R}^{M \times K}$ and $\mathbf{a}_K \in \mathbb{R}^{M \times 1}$ in equation (1.1).

The operators used throughout the text are organized in Table 1.1.

Table 1.1: Operators used throughout this work

| **Operator** | **Input** | **Output** |
|---|---|---|
| $(\cdot)^{\mathrm{T}}$ | vector or matrix | input vector or matrix with transposed elements |
| $(\cdot)^{\mathrm{H}}$ | vector or matrix | input vector or matrix with transposed and conjugated elements |
| $(\cdot)^{*}$ | vector or matrix | vector or matrix of the complex conjugate elements of the input vector or matrix |
| $(\cdot)^{-1}$ | matrix | inverse of input matrix |
| $\|\mathbf{x}\|_p$ | vector | $p$-norm, $\left(\sum_{m=1}^{M} \|x_m\|^p\right)^{1/p}$ |
| $\|\cdot\|_0$ | vector | number of non-zero entries of input vector |
| $\mathrm{diag}\,(\cdot)$ | vector | diagonal matrix where the diagonal entries are the elements of the input vector |
| $\det(\cdot)$ | square matrix | determinant of the input matrix |
| $\mathrm{tr}(\cdot)$ | square matrix | trace of the matrix, that is, the sum of the diagonal elements of the input matrix |
| $\langle\cdot,\cdot\rangle$ | two vectors | inner product between two input vectors |
| $\mathrm{vec}(\cdot)$ | matrix | column vector which is obtained by transposing the rows of the input matrix and stacking them up |
| $\mathrm{rem}(\cdot)$ | matrix | matrix with zero columns removed |

# Chapter 2

# Basic Concepts of Cellular MIMO Systems

In this chapter, we describe the principles of cellular networks and the improvements achieved by using the multiple-input multiple-output approach. Furthermore, we introduce the main concepts required to understand the following chapters.

## 2.1 Cellular Networks

In cellular or more general mobile networks, the users' terminals in a given geographic area are served by several BSs [24]. Each BS serves simultaneously a certain number of terminals located in the coverage area of the BS, as illustrated in Figure 2.1. Such a coverage area is called a *cell*, allowing an extensive geographic area to be partitioned into *cells* [25]. The communication between the terminals and BS is two-way or in duplex format.

In the *Downlink* (DL), the BS transmits signals to the assigned terminals, whereas in the *Uplink* (UL), the terminals transmit signals to the BS, as depicted in Figure 2.1. In general, time-division duplex and frequency-division duplex are used as duplex transmission schemes, which are discussed in subsection 2.1.1.

The cellular networks were originally conceived for voice communication. However, nowadays the majority of traffic in wireless networks accounts for viewing of video content.

A relevant performance metric of cellular networks is the *area throughput*. The area throughput performance metric, also discussed in [26], is given by

$$\text{Area throughput}[\text{bit/s/km}^2] = B[\text{Hz}] \cdot D[\text{cells/km}^2] \cdot \text{SE}[\text{bit/s/Hz/cell}] \qquad (2.1)$$

where $B$ is the communication bandwidth, $D$ is the average cell density and SE is

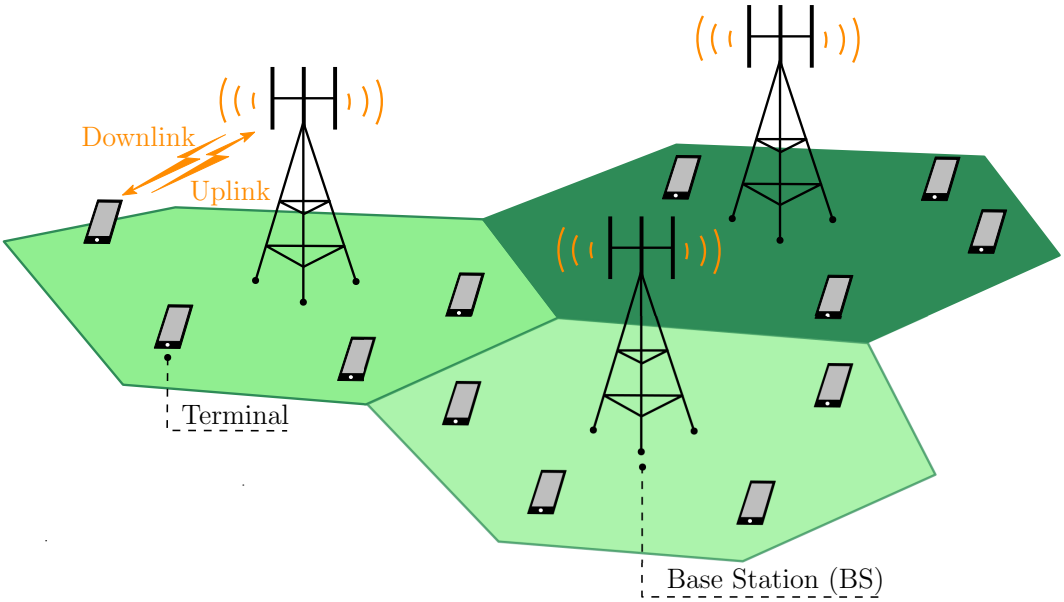



Figure 2.1: A basic cellular network, where each BS provides service to all terminals in the cell.

the spectral efficiency per cell. Therefore, to improve the area throughput of cellular networks, we can

- allocate more bandwidth;
- add more BSs;
- improve the SE.

Increasing the bandwidth is not so attractive since the frequency spectrum is a natural resource with high cost and facing scarcity. Furthermore, it also entails using much higher frequency bands which limits the communication range. Despite that, millimeter wavelength (mmWave) bands (e.g., in the range 30–300 GHz) can be used for short-range applications [27]. Densifying the network, by employing more BSs, is a hard task without moving BSs closer to terminals, which leads to increased risks of being in deep shadow, thereby reducing coverage.

These are the main reasons why it is preferable to enhance the SE. One way to improve SE is by employing multiple antennas at the BS, to collect more energy from the electromagnetic waves [26]. This promising solution is detailed in section 2.2.

### 2.1.1 Duplexing schemes

For cellular systems, the possibility of transmitting and receiving data in both directions using the same environment is essential. Then, choosing the proper duplex

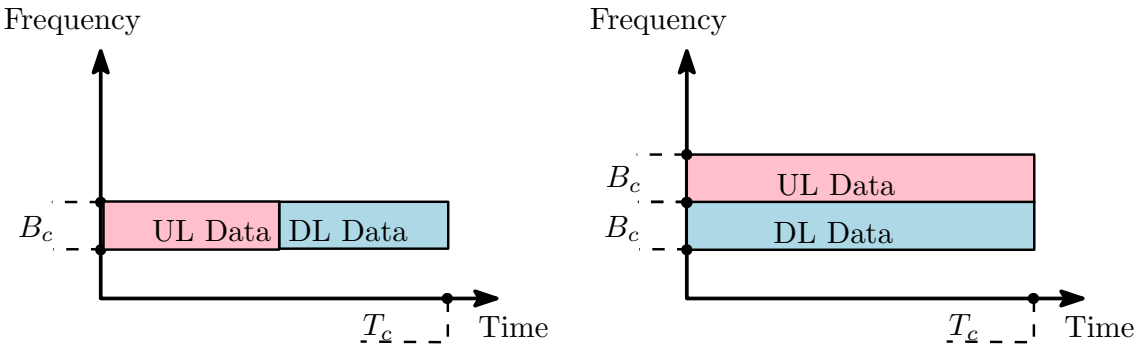
scheme is a fundamental part of the overall specification for the cellular system. The main duplex schemes used in wireless communications are time-division duplex and frequency-division duplex, which are illustrated in Figure 2.2.

### Time-Division Duplex (TDD)

In TDD, in a single frequency, the transmission and reception of data are periodically alternated [28]. The uplink and downlink directions are separated in the time domain as one can see in Figure 2.2a. As the same band is used, uplink and downlink channel responses are reciprocal to each other due to the RF channel reciprocity between the transmitter and receiver during short period of times [29]. Then, once we have a channel estimate of the uplink direction at the transmitter, we can directly utilize it as an estimate of the downlink channel.

### Frequency-Division Duplex (FDD)

As the name suggests, in the FDD scheme, the uplink and downlink are separated in the frequency domain [28]. When FDD is used, it is possible to transmit and receive signals simultaneously as the frequency channels, depicted in Figure 2.2b, used by the receiver and transmitter are different. Since uplink and downlink are on different frequency bands that might be far apart, the channels are not reciprocal [29]. Thereby, the receiver communicates the estimated channel state information (CSI) over a control feedback channel which can be very costly as the number of antennas at the transmitter increases.




(a) TDD. (b) FDD.

Figure 2.2: Duplexing Schemes.

## 2.1.2 Coherence Interval

The time during which the channel variation in time is negligible is called *coherence time* and $T_c$ is used to denote it. In addition, the length of a frequency interval over which the channel frequency response is approximately flat is called *coherence*

*bandwidth*, $B_c$. Then, a time-frequency space of duration $T_c$ and bandwidth $B_c$ is called *coherence interval* and it is the largest possible time-frequency space in which the effect of the channel can be reduced to a multiplication by a scalar gain. The coherence interval has the length [26]

$$\tau_{\mathrm{c}} = B_c T_c \ \text{ samples.} \tag{2.2}$$

The transmission of samples in Figure 2.2 is performed during the coherence interval. For the TDD mode, for example, the coherence interval can be structured as in Figure 2.3, in which $\tau_{\mathrm{ul}}$ and $\tau_{\mathrm{dl}}$ is the number of samples per coherence interval allocated for transmission of uplink and downlink data, respectively. Moreover, $\tau_{\mathrm{p}}$ is the number of samples reserved for pilot symbols in the uplink.

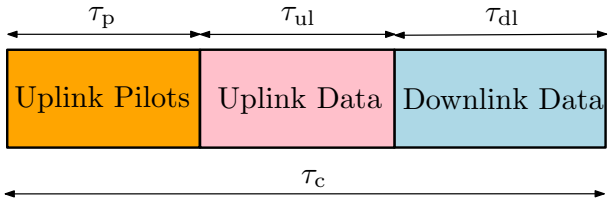



Figure 2.3: Allocation of the samples in a coherence interval (TDD mode).

## 2.2 MIMO

The concept of using multiple antennas was first suggested in 1910, for transoceanic communications [30]. Since then, many studies have been demonstrating how multiple-input multiple-output (MIMO) is capable of enhancing the quality and capacity of wireless communications [31–40].

The MIMO performance depends on the propagation conditions. NLOS (Non-line-of-sight) and LOS (line-of-sight) are the most common classification. When a strong direct path is available from the transmitter to the receiver, we have LOS propagation condition as illustrated in Figure 2.4, whereas multiple obstructed paths are predominant in NLOS propagation condition. This multipath propagation that generally impairs the performance of single-antenna systems is exploited in MIMO to provide a robust link [24]. This is possible since the received signals from one transmitting antenna may differ from the others in phase, timing, and signal strength characteristics. On the other hand, LOS condition leads to a high receive signal-to-noise ratio SNR which will also contribute to improve MIMO capacity. A performance comparison between LOS and NLOS conditions is addressed in [41]. Nevertheless, obstacles are so frequent in urban and indoor environments that clear LOS between transmitter and receiver is very rare. In general, the MIMO technique can be configured as single-user MIMO, multi-user MIMO, and massive MIMO.

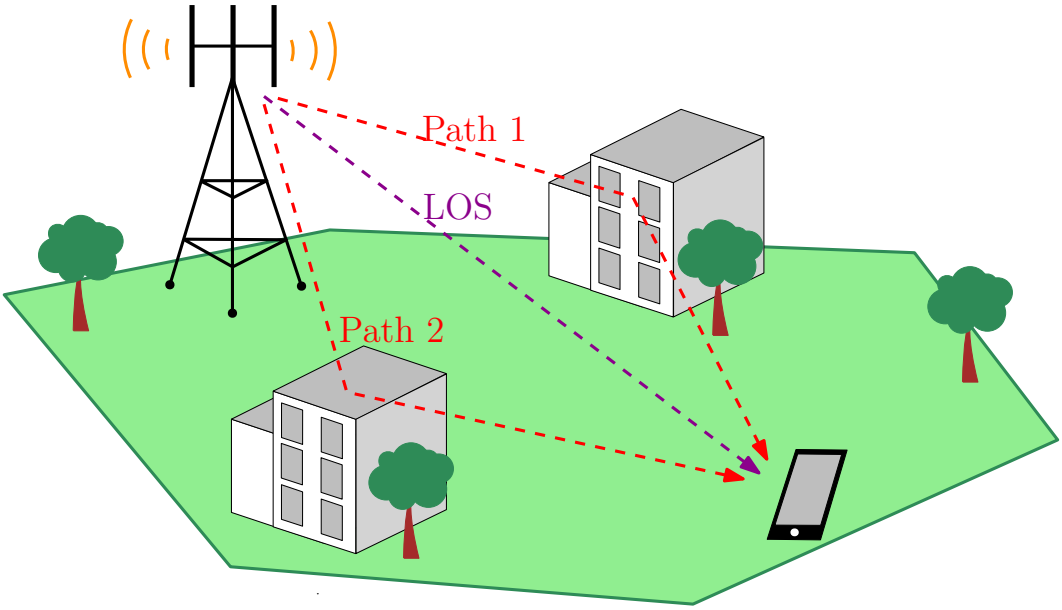



Figure 2.4: Multipath propagation in a single-cell.

### 2.2.1 Single-user MIMO

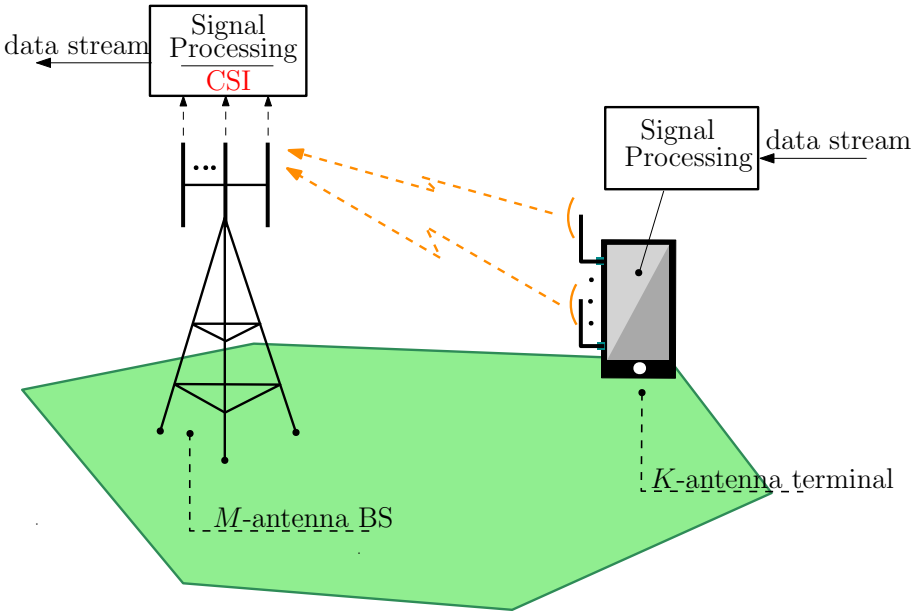
SU MIMO or point-to-point MIMO consists of a transmitter BS and a terminal receiver, both equipped with an array of antennas [18, 41]. An illustration is provided in Figure 2.5, where the $M$-antenna BS serves a certain $K$-antenna terminal. In SU MIMO, the BS transmits multiple streams to the same terminal and different terminals are orthogonally multiplexed [42]. In theory, under rich scattering and time-invariant environments MIMO systems can provide remarkable capacity growth [40, 43, 44]. Shannon theory yields the following capacity formulas (in b/s/Hz) for uplink and downlink

$$C_{\text{ul}} = \log\det\left(\mathbf{I}_{(M)} + \frac{\rho_{\text{ul}}}{K}\mathbf{G}\mathbf{G}^{\text{H}}\right), \tag{2.3}$$

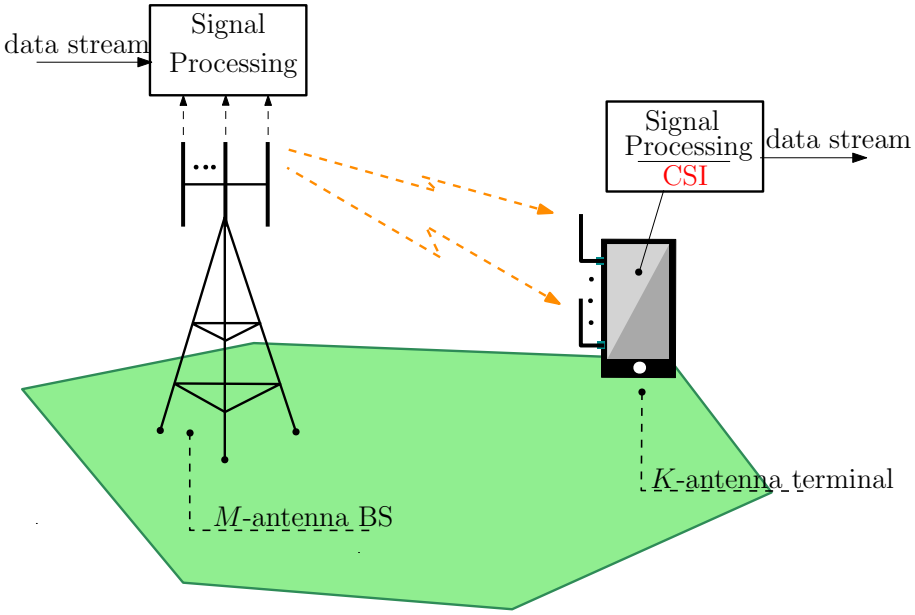
$$C_{\text{dl}} = \log\det\left(\mathbf{I}_{(M)} + \frac{\rho_{\text{dl}}}{M}\mathbf{G}\mathbf{G}^{\text{H}}\right). \tag{2.4}$$

when there is additive white Gaussian noise (AWGN) at the receiver. The uplink and downlink SNRs are denoted by $\rho_{\text{ul}}$ and $\rho_{\text{dl}}$, and $\mathbf{G}$ is the matrix comprising all the channel gains between the BS and the terminal. However, equations (2.3) and (2.4) are just theoretical bounds, based on assumptions about the underlying time-varying channel model and perfect CSI at the receiver as well as at the transmitter [18]. CSI is optional at BS, but is mandatory at the terminal which complicates the terminal equipment. The terminal also requires separate Radio Frequency (RF) chains per antenna, where an RF chain includes filters, low-noise amplifiers (LNAs), and down-conversion mixers [45, 46]. Moreover, the terminal demands advanced signal processing to separate the data streams, which makes the terminal device even

more complex. In addition, the propagation environment must support $\min(M, K)$ data streams, which is difficult to achieve in practice when compact array antennas are used. Another disadvantage of SU MIMO is that when the SNR is low, i.e., at the cell edge, the spectral efficiency is slowly improved as the number of BS antennas increases. Thus, the relevance of SU MIMO is limited in practice.




(a) Uplink.




(b) Downlink.

Figure 2.5: Single-user MIMO.

### 2.2.2 Multi-user MIMO

By splitting the antenna array at the terminal receiver of a SU MIMO system into autonomous antennas, we obtain the multi-user (MU) MIMO. In MU MIMO, a BS equipped with $M$ antennas transmits simultaneously spatially multiplexed streams to $K$ different terminal users [42]. Figure 2.6 depicts a MU MIMO system in both uplink and downlink. MU MIMO allows the spatial multiplexing gain at the BS to be accomplished with single-antenna terminals [47]. This enables the development of small and low-cost terminals, whereas the higher complexity is kept at the BS. Nevertheless, using multiple antennas at the terminal can be viewed as optional equipment allowing extra diversity gain and enhancing the received signal quality.

MU MIMO differs from SU MIMO in some aspects. First, the terminals are typically separated by many wavelengths. In this way, MU MIMO overcomes most of propagation limitations in MIMO such as ill-behaved channels due to the small spacing between the receiving antennas. Second, in MU MIMO, the terminals cannot collaborate among themselves, either to transmit or to receive data [41, 47, 48]. The uplink and downlink capacity bounds are given by [18, 49]

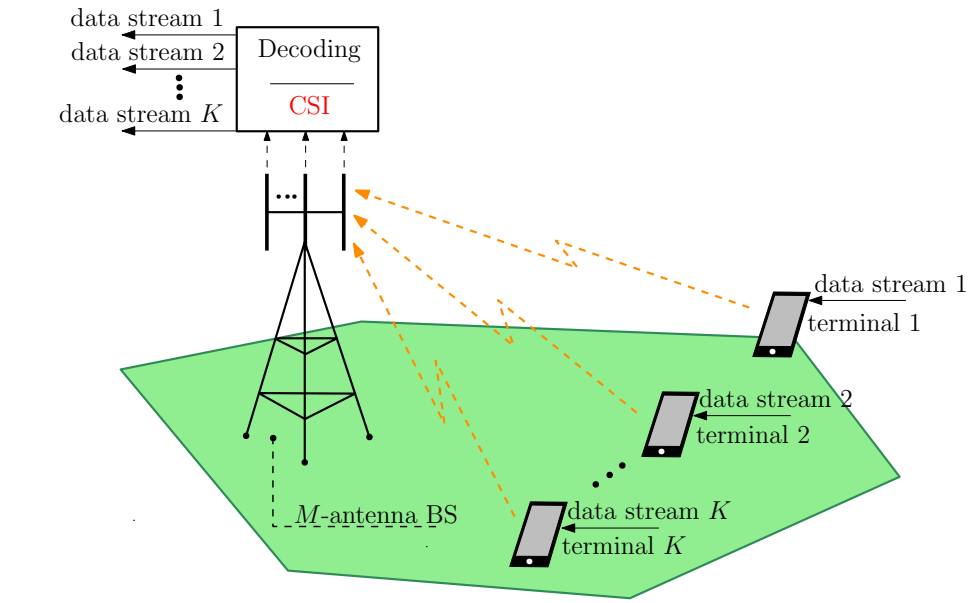
$$C_{\text{ul}} = \log\det\left(\mathbf{I}_{(M)} + \rho_{\text{ul}}\mathbf{G}\mathbf{G}^{\text{H}}\right), \tag{2.5}$$

$$C_{\text{dl}} = \max_{\substack{\eta_k \geq 0 \\ \sum_{k=1}^{K} \eta_k \leq 1}} \log\det\left(\mathbf{I}_{(M)} + \rho_{\text{dl}}\mathbf{G}\text{diag}\left(\boldsymbol{\eta}\right)\mathbf{G}^{\text{H}}\right) \tag{2.6}$$

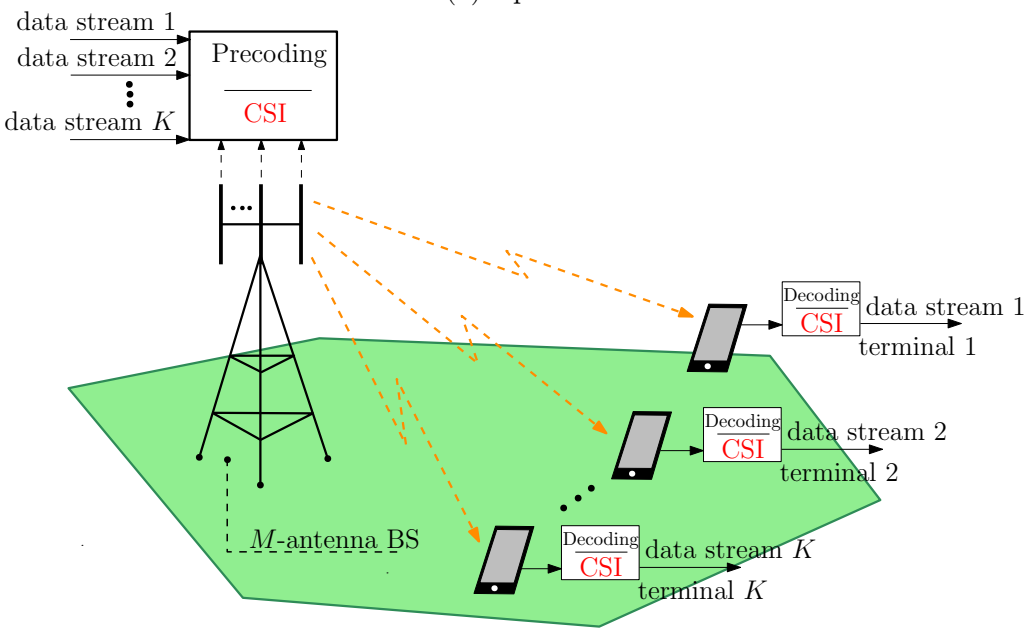
where $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_K]$ is the vector comprising the power allocated for each user. The computation of downlink capacity according to (2.6) requires the solution of a convex optimization problem. In fact, it is a power allocation problem and can be solved with iterative water-filling algorithms [50, 51].

As the terminals are served by the BS in the same time-frequency resource, strong interference among the users is present. When the signal is precoded at the BS, the interference among the users is generally reduced. Therefore, the BS must know the CSI perfectly in comparison with SU MIMO schemes, since CSI is of critical importance to most precoding techniques. Precoding strategies include linear and nonlinear approaches, which will be discussed in chapter 3.

Unfortunately, complicated signal processing by both the BS and the terminals are required to achieve the capacity bounds in equations (2.5) and (2.6). Furthermore, in the downlink, both the BS and the terminals must have CSI, which requires substantial resources to be set aside for transmission of pilots in both directions. Thus, the implementation of MU MIMO is also limited in practice.




(a) Uplink.



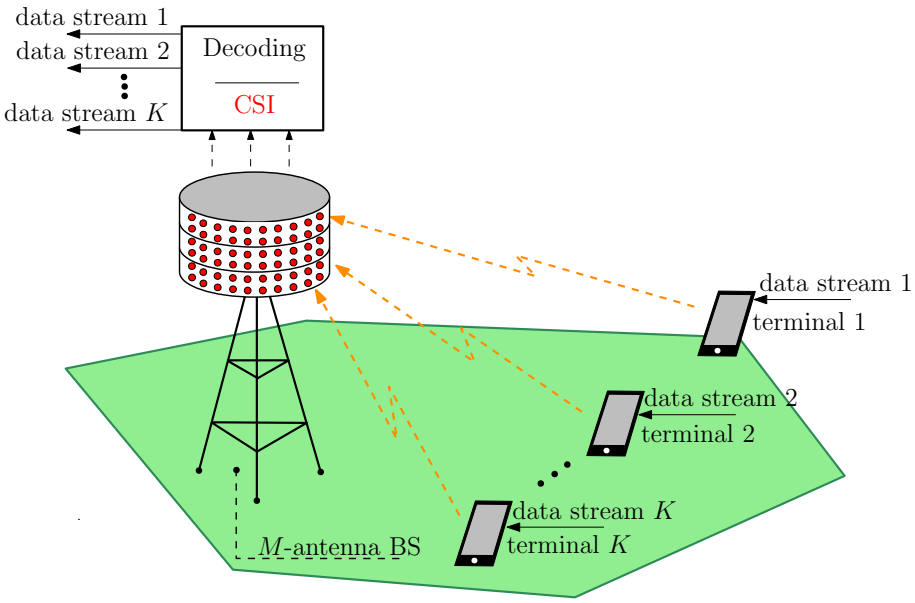

(b) Downlink.

Figure 2.6: Multi-user MIMO.

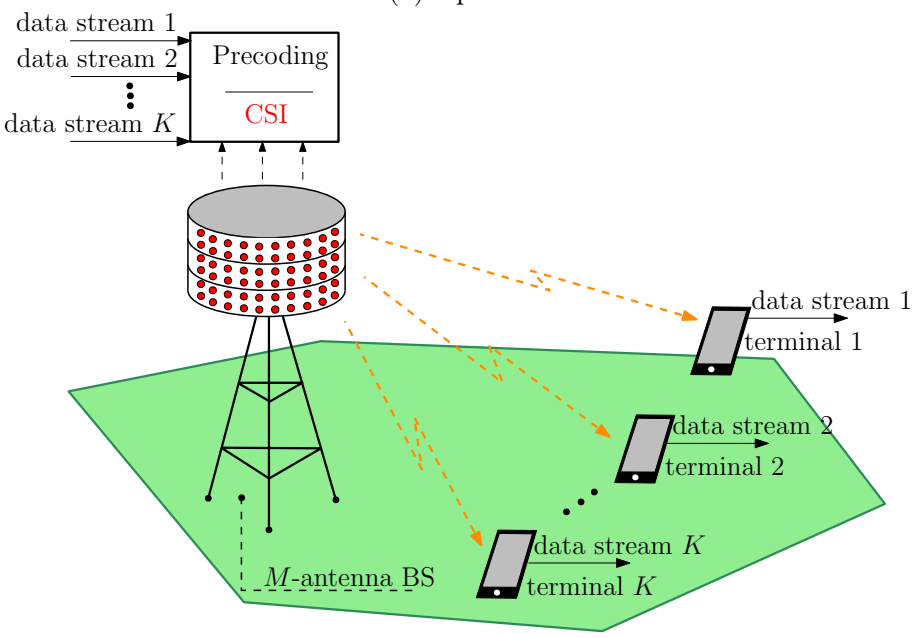
### 2.2.3 Massive MIMO

The massive MIMO concept was introduced by Marzetta in [52], where the main idea was to scale up MU MIMO systems, by deploying a huge number of antennas at the BS that dramatically exceeds the number of active terminals, $M \gg K$ . Thereby, the use of simple linear processing becomes nearly optimal. As in MU MIMO, the base station serves all active users, simultaneously, in the same time-frequency resource. Figure 2.7 illustrates a massive MIMO system operating in both uplink and downlink modes. Since the number of BS antennas is large, the

channel becomes nearly deterministic, and hence, the effects of small-scale fading and frequency dependence disappear. This is known as *channel hardening*. One can compare Figures 2.7 and 2.6 and see that the main difference between MU MIMO and massive MIMO occurs in the downlink, where the terminals no longer need to perform decoding. Thus, no channel estimation is required at the terminals and only the BS obtains CSI, thanks to channel hardening. By operating in TDD mode and exploiting reciprocity of the propagation channel, the amount of resources needed for pilots only depends on the number of simultaneously served terminals, $K$.




(a) Uplink.




(b) Downlink.

Figure 2.7: Massive MIMO.

## 2.3 Concluding Remarks

In this chapter, the cellular networks were introduced. A *cell* was described as the geographic area covered by a transmit facility called base station. The users or terminals in a cell communicate with the base station via uplink or downlink. Among the aspects that compose the presented performance metric known as area throughput, the spectral efficiency was shown to be preferable. The MIMO technique was suggested as a promising solution to improve the spectral efficiency. In addition, several MIMO configurations were presented culminating in the target: massive MIMO. In the next chapter, the details of the massive MIMO systems are discussed, regarding the system model, propagation and also some challenges that are inherent to this technique.

# Chapter 3

# Key Features of Massive MIMO

When the number of BS antennas becomes significantly larger than the number of terminal users in the cell, we are in the context of large or massive MIMO. The benefits of massive MIMO include improvements in throughput and radiated energy efficiency [17], interference suppression among users [26], and more effective linear signal processing at BS [53]. Nevertheless, the advantages are followed by some challenges like pilot contamination and unfavorable propagation. In this chapter, we provide an overview of massive MIMO theory, including some preliminaries and essential concepts. Both uplink and downlink system designs are presented. Furthermore, the massive MIMO challenges are discussed. For the sake of simplicity, we restrict our discussion to the single-cell systems, and hence we consider interference from neighboring cells to be negligible.

## 3.1 System Model

We consider a single-cell massive MIMO system where a BS equipped with $M$ antennas simultaneously serves $K$ single-antenna terminal users. The massive MIMO operating in both uplink and downlink is illustrated in Figure 2.7. The majority of massive MIMO deployment rely on the TDD transmission scheme [17, 18, 26]. As discussed in subsection 2.2, TDD is useful due to the effective acquisition of downlink CSI at the BS by exploiting channel reciprocity. On the other hand, vast amounts of spectrum are reserved for FDD operation [54] and considerably reduced CSI accuracy can result [55, 56]. However, the large number of BS antennas leads to a CSI overhead increase when operating in FDD mode. Therefore, TDD is preferable for massive MIMO [57], and it is the scheme considered in this work.

Now we need to model the channel between the BS and the terminal. In fact, it depends on the propagation environment, which is generally classified as LOS, NLOS with rich scattering and NLOS with spatial multipath [26]. The uplink and downlink system models are defined in subsections 3.1.2 and 3.1.3.

### 3.1.1 Channel Model

In a single-cell system, the channel response for terminal $k$ to the BS is denoted by $\mathbf{g}_k \in \mathbb{C}^{M\times 1}$, where each element of $\mathbf{g}_k$ is the channel response from the terminal to one of the $M$-BS antennas. Let $\mathbf{G}$ be a matrix comprising the channel responses for all the terminals in the cell to the BS,

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1K} \\ \vdots & \vdots & \ddots & \vdots \\ g_{M1} & g_{M2} & \cdots & g_{MK} \end{bmatrix}, \tag{3.1}$$

and $\mathbf{g}_k$ is the $k$th column of $\mathbf{G}$. An illustration of the channel matrix is provided in Figure 3.1. The model choice for the elements of $\mathbf{G}$ depends on the propagation environment, generally classified as LOS and NLOS, both mentioned in section 2.2.

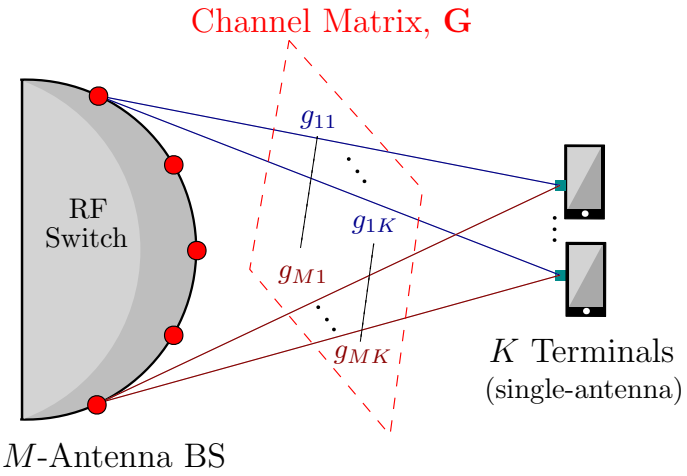



Figure 3.1: Channel matrix between BS and terminals.

#### LOS channel model

When there is no local scattering, and all terminals have a direct path to the BS, we are in the LOS case which is depicted in Figure 3.2. Consider that the BS is a horizontal uniform linear array (ULA) with antenna spacing $d_H$, which is measured in the number of wavelengths between adjacent antennas. Moreover, assume that the terminals are located at fixed locations in the far-field of the BS array. Hence, we have the following deterministic channel response between the BS and a terminal $k$ located at the angle $\theta_k \in [0, 2\pi)$

$$\mathbf{g}_k = \sqrt{\beta_k}\left[1 \;\; e^{2\pi j d_H \sin(\theta_k)} \ldots e^{2\pi j d_H (M-1)\sin(\theta_k)}\right]^T, \tag{3.2}$$

where $\beta_k$ describes the macroscopic large-scale fading [58]. Observe that the phase rotations in (3.2) of adjacent antennas differ by $d_H \sin(\theta_k)$, which is the additional

distance that signals have to travel, as illustrated in Figure 3.2. The large-scale coefficient is often modeled in decibels as

$$\beta_k = \Gamma - 10\alpha \log_{10}\left(\frac{d_k}{1\text{km}}\right) + f_k \tag{3.3}$$

where $d_k$(km) is the distance between the BS and the terminal, $\alpha$ is the path-loss exponent which determines how fast the signal power decays with the distance, and $\Gamma$ determines the median channel gain at a reference distance of 1 km [26]. Deterministic parameters $\Gamma$ and $\alpha$ are functions of the carrier frequency, antenna gains, and vertical heights of the antennas which can be obtained by established propagation models [59]. The term $f_k$ is the realization of a random variable $F_k \sim (0, \sigma_{\text{sf}}^2)$ which models the *shadow fading* effect [60]. The shadow fading can be viewed as a model of physical blockage from large obstacles.

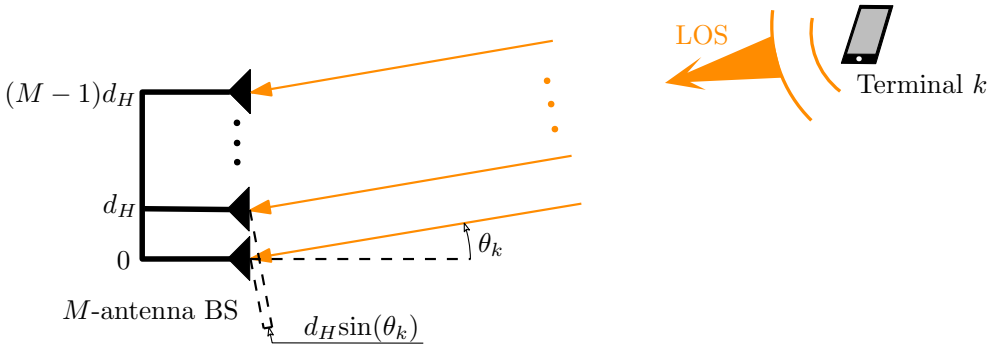



Figure 3.2: LOS propagation between terminal $k$ and a BS equipped with a ULA with $M$ antennas.

### NLOS channel model

In the NLOS case, there are two possibilities for the scattering scenarios. In the first, the system operates in a rich scattering environment so that the signal arises at BS through many paths. As illustrated in Figure 3.3, a single-antenna terminal communicates with the BS equipped with a uniform linear array of $M$ antennas. The LOS path is blocked, but the signal finds multiple other paths via specular reflection. We can use the *uncorrelated Rayleigh fading* also known as independent and identically distributed (i.i.d.) Rayleigh fading [26] to model $\mathbf{g}_k$ as a realization of the random variable

$$\boldsymbol{g}_k \sim \mathcal{CN}(\mathbf{0}_{(M)},\, \beta_k \mathbf{I}_{(M)}). \tag{3.4}$$

Thus, $g_{mk}$ can be modeled as

$$g_{mk} = \sqrt{\beta_k} h_{mk} \tag{3.5}$$

where $\beta_k$ is a large-scale coefficient dependent only on $k$ and $h_{mk}$ is the realization of a random variable distributed as $\mathcal{CN}(0,\,1)$ which represents the effect of small-scale fading.

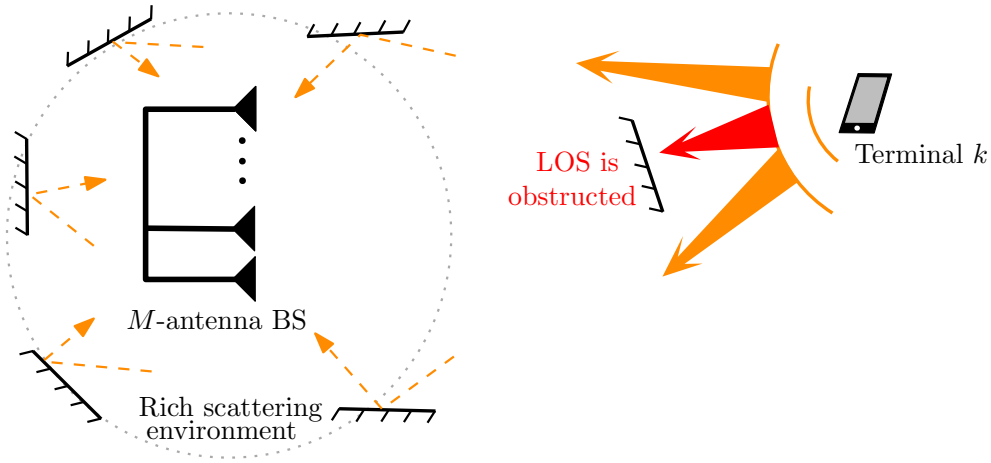



Figure 3.3: NLOS propagation with uncorrelated Rayleigh fading between the ULA-BS and the terminal $k$.

In the second case, the BS receives the signal only through a few number of significant paths. It can happen due to insufficient antenna separation or poor scattering environment [61]. Hence, the received signal has a strong spatial signature in the sense that stronger average signal gains are received from certain spatial directions [26]. In this way, the channel can be modeled by *correlated Rayleigh fading* [26] in which $\mathbf{g}_k$ is modeled as a realization of the random variable

$$\boldsymbol{g}_k \sim \mathcal{CN}(\mathbf{0}_{(M)},\, \mathbf{R}_k). \tag{3.6}$$

where $\mathbf{R}_k \in \mathbb{C}^{M\times M}$ is the spatial correlation matrix and it is also the covariance matrix as the mean is zero. Similarly to the uncorrelated Rayleigh fading model, the Gaussian distribution is used to model the small-scale fading variations and the covariance matrix describes the large-scale fading. In the uncorrelated fading the covariance matrix is diagonal, whereas in the correlated fading the covariance matrix is not necessarily diagonal. A possible model for $\mathbf{R}_k$ is described in [26, 35]. This scenario is exemplified in Figure 3.4, where the signal arrives at BS through only two paths. The received signal at the BS is the superposition of $N$ multipath components. Each of the multipath components thus results in a plane wave that reaches the array from a particular angle $\theta_{kn}$ and gives an *array response* or *steering vector* $\mathbf{a}_{kn} \in \mathbb{C}^{M\times 1}$ similar to the LOS case in (3.2)

$$\mathbf{a}_{kn} = h_{kn}\left[1 \;\; e^{2\pi j d_H \sin(\theta_{kn})} \ldots e^{2\pi j d_H (M-1)\sin(\theta_{kn})}\right]^{\mathrm{T}} \tag{3.7}$$

where $h_{kn} \in \mathbb{C}$ accounts for the gain and phase-rotation for this path. Then, the

channel response for terminal $k$ is the superposition

$$\mathbf{g}_k = \sum_{n=1}^{N} \mathbf{a}_{kn} \tag{3.8}$$

of the array responses of the $N$ components.

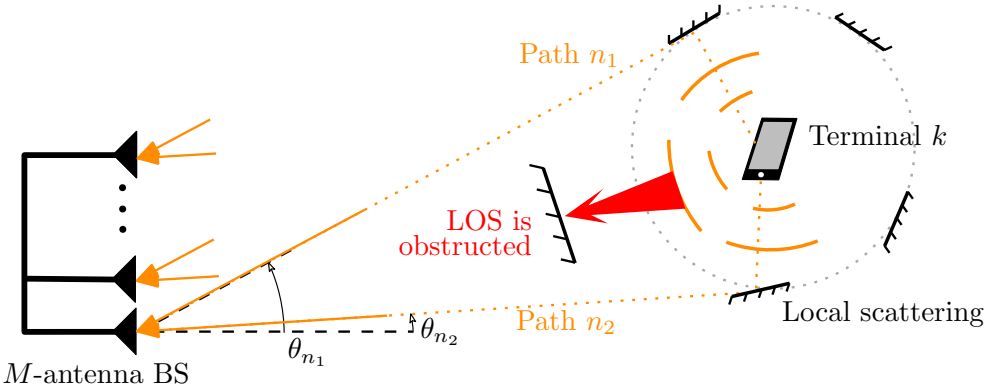



Figure 3.4: NLOS propagation with correlated Rayleigh fading between the ULA-BS and the terminal $k$.

Suppose the angles $\theta_{kn}$ are realizations of i.i.d. random variables with angular probability density function (PDF) $f(\theta)$ and $h_{kn}$ are i.i.d. random variables with zero-mean and variance $E[|h_{kn}|^2]$. The variance represents the average gain of the $n$th path and the total average gain of the multipath components is denoted by $\beta_k = \sum_{n=1}^{N} E[|h_{kn}|^2]$. If the number of paths is large, the multidimensional central-limit theorem [62] implies that

$$\boldsymbol{g}_k \to \mathcal{CN}(\mathbf{0}_{(M)}, \mathbf{R}_k), \quad N \to \infty \tag{3.9}$$

where the convergence is in distribution and the correlation matrix is

$$\mathbf{R}_k = E\left[\sum_{n=1}^{N} \boldsymbol{a}_n \boldsymbol{a}_n^{\mathrm{H}}\right]. \tag{3.10}$$

### 3.1.2 Uplink

The signal $\mathbf{y} \in \mathbb{C}^{M\times 1}$ received at BS equipped with $M$ antennas is of the form

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{g}_k q_k + \mathbf{w} = \mathbf{G}\mathbf{q} + \mathbf{w} \tag{3.11}$$

where $\mathbf{w}$ is a realization of a random variable distributed as $\mathcal{CN}(\mathbf{0}_{(M)}, \sigma_w^2 \mathbf{I}_{(M)})$ and represents independent additive receiver noise with zero mean and variance $\sigma_w^2$. The signal $\mathbf{q} \in \mathbb{C}^{K\times 1}$ is the composition of the symbols transmitted from all the terminals

in the cell. The element $q_k$ is the symbol transmitted by terminal $k$ and can be either a data or pilot signal. The channels are constant within a coherence block, while the signals and noise take new realization at every sample. When the terminals are transmitting data symbols, the BS performs linear detection to separate the symbols from each terminal. On the other hand, when the terminals transmit pilot symbols, the BS uses them to perform channel estimation. The channel estimation process is detailed in section 3.2, whereas the data reception is outlined in section 3.4.

### 3.1.3 Downlink

The received signal $y_k$ at terminal $k$ is modeled as

$$y_k = \mathbf{g}_k^{\mathrm{T}}\mathbf{x} + w_k = \mathbf{g}_k^{\mathrm{T}}\mathbf{P}\mathbf{q} + w_k \tag{3.12}$$

in which $w_k$ is a realization of a random variable distributed as $\mathcal{CN}(0, \sigma_w^2)$ representing the receiver noise. The channel between terminal $k$ and BS is denoted by $\mathbf{g}_k \in \mathbb{C}^{M\times 1}$. The BS transmits the vector $\mathbf{x} \in \mathbb{C}^{M\times 1}$ which is a combination of the intended vector $\mathbf{q}$ and the precoding matrix $\mathbf{P}$. The precoding operation is explained in subsection 3.3.1.

## 3.2 Channel Estimation in the Uplink

The massive number of antennas plays a significant role in the estimation of the channel gains between the BS and the terminals. Indeed, activating more BS antennas enables the BS to attain an accurate estimation of the channels so that the chance of success in the precoding is increased.

In each coherence block, $\tau_{\mathrm{p}}$ of the available $\tau_{\mathrm{c}}$ samples are reserved for the terminals to transmit the pilot symbols. The terminals simultaneously transmit their corresponding pilots, which are known at the BS. As the $K$ pilot waveforms share the same coherence block, pilots from different users need to be mutually orthogonal to avoid interference with each other. The pilot sequence assigned to terminal $k$ is denoted by $\boldsymbol{\phi}_k \in \mathbb{C}^{\tau_{\mathrm{p}}\times 1}, \tau_{\mathrm{p}} \geq K$ and is assumed to have unit power magnitude, $\|\boldsymbol{\phi}_k\|_2^2 = 1$. We also denote the pilot matrix as $\boldsymbol{\Phi} \in \mathbb{C}^{\tau_{\mathrm{p}}\times K}$, which contains the pilot sequences from all terminals, such that

$$\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi} = \mathbf{I}_{(K)}. \tag{3.13}$$

To set the total energy spent by each terminal equal to the duration of the pilot sequence $\tau_{\mathrm{p}}$, the pilot sequences are scaled by $\sqrt{\tau_{\mathrm{p}}}$ [18]. Hence, complete set of pilot

signals transmitted by the terminals is given by

$$\mathbf{X}_{\mathrm{p}} = \sqrt{\tau_{\mathrm{p}}}\boldsymbol{\Phi}^{\mathrm{H}}. \tag{3.14}$$

The BS receives the $M \times \tau_{\mathrm{p}}$ signal

$$\begin{aligned} \mathbf{Y}_{\mathrm{p}} &= \mathbf{G}\mathbf{X}_{\mathrm{p}} + \mathbf{W}_{\mathrm{p}} \\ &= \sqrt{\tau_{\mathrm{p}}}\mathbf{G}\boldsymbol{\Phi}^{\mathrm{H}} + \mathbf{W}_{\mathrm{p}}. \end{aligned} \tag{3.15}$$

in which the entries of the $M \times \tau_{\mathrm{p}}$ noise matrix $\mathbf{W}_{\mathrm{p}}$ are independent identically distributed (i.i.d) realizations of a random variable distributed as $\mathcal{CN}(0,1)$.[1]

To estimate $\mathbf{g}_k$, BS can correlate $\mathbf{Y}_{\mathrm{p}}$ in equation (3.15) with the pilot sequence $\boldsymbol{\phi}_k$. In this way, the received pilot signal is

$$\begin{aligned} \mathbf{y}'_{\mathrm{p}_k} &= \mathbf{Y}_{\mathrm{p}}\boldsymbol{\phi}_k \\ &= \sqrt{\tau_{\mathrm{p}}}\mathbf{g}_k\boldsymbol{\phi}_k^{\mathrm{H}}\boldsymbol{\phi}_k + \sum_{\substack{l=1,\\ l\neq k}}^{K} \sqrt{\tau_{\mathrm{p}}}\mathbf{g}_l\boldsymbol{\phi}_l^{\mathrm{H}}\boldsymbol{\phi}_k + \mathbf{W}_{\mathrm{p}}\boldsymbol{\phi}_k. \end{aligned} \tag{3.16}$$

Since we are considering orthogonal pilots, the inner product $\boldsymbol{\phi}_l^{\mathrm{H}}\boldsymbol{\phi}_k = 0$ and thus we can obtain the noisy version of $\mathbf{g}_k$,

$$\mathbf{y}'_{\mathrm{p}_k} = \sqrt{\tau_{\mathrm{p}}}\mathbf{g}_k + \mathbf{W}_{\mathrm{p}}\boldsymbol{\phi}_k. \tag{3.17}$$

Equivalently, the BS can estimate the channel matrix $\mathbf{G}$ by right-multiplying the equation (3.15) by the pilot matrix. Then, yielding

$$\begin{aligned} \mathbf{Y}'_{\mathrm{p}} &= \mathbf{Y}_{\mathrm{p}}\boldsymbol{\Phi} \\ &= \sqrt{\tau_{\mathrm{p}}}\mathbf{G}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi} + \mathbf{W}\boldsymbol{\Phi} \\ &= \sqrt{\tau_{\mathrm{p}}}\mathbf{G} + \mathbf{W}\boldsymbol{\Phi} \end{aligned} \tag{3.18}$$

which is a noisy version of the channel matrix $\mathbf{G}$.

### 3.2.1 Pilot sequence generation

The pilot matrix can be generated as a Walsh-Hadamard matrix $\boldsymbol{\Phi}_{(\tau_{\mathrm{p}})}$ [26]. The Walsh-Hadamard of dimension $2^N$ can be computed with the following recursive formula

$$\boldsymbol{\Phi}_{(2^N)} = \begin{bmatrix} \boldsymbol{\Phi}_{(2^{N-1})} & \boldsymbol{\Phi}_{(2^{N-1})} \\ \boldsymbol{\Phi}_{(2^{N-1})} & -\boldsymbol{\Phi}_{(2^{N-1})} \end{bmatrix} = \boldsymbol{\Phi}_{(2)} \otimes \boldsymbol{\Phi}_{(2^{N-1})}, \tag{3.19}$$

[1] For convention, the noise variance is equal to one

in which $\otimes$ denotes the Kronecker product. The initial seed is $\mathbf{\Phi}_{(0)} = 1$ and hence,

$$\mathbf{\Phi}_{(2^1)} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad \mathbf{\Phi}_{(2^2)} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{3.20}$$

are examples of Walsh-Hadamard matrices of dimensions 2 and 4, respectively.

### 3.2.2 MMSE Channel Estimation

If we have a *priori* knowledge of the signal distributions, a Bayesian estimator can be employed to obtain a better version of the estimated channel matrix $\mathbf{G}$. Assume we want to estimate a vector $\mathbf{x} \in \mathbb{C}^{M\times 1}$ which is a realization of a Gaussian random variable with zero mean and covariance $\mathbf{R}$, from the observation $\mathbf{y} = \mathbf{A}\mathbf{x}+\mathbf{w} \in \mathbb{C}^{K\times 1}$. The matrix $\mathbf{A}$ is considered known and $\mathbf{w}$ is additive white Gaussian noise (AWGN) with zero mean and variance $\sigma_w^2$. As defined in [26], the MMSE estimator of $\mathbf{x}$ based on $\mathbf{y}$ is

$$\hat{\mathbf{x}}_{\text{MMSE}} = \mathbf{R}\mathbf{A}^{\text{H}}(\sigma_w^2\mathbf{I} + \mathbf{A}\mathbf{R}\mathbf{A}^{\text{H}})^{-1}\mathbf{y}. \tag{3.21}$$

We assume the large-scale fading coefficients are known, then $g_{mk}$ is a realization of a random variable with known prior distribution of $\mathcal{CN}(0,\, \beta_k)$. By using equation (3.21) we can obtain the MMSE estimator of $\mathbf{g}_k$ as

$$\widehat{\mathbf{g}}_k = \frac{\sqrt{\tau_{\text{p}}}\beta_k}{1 + \tau_{\text{p}}\beta_k}\mathbf{y}'_{\text{p}_k}. \tag{3.22}$$

The complete channel estimation procedure is detailed in Figure 3.5.

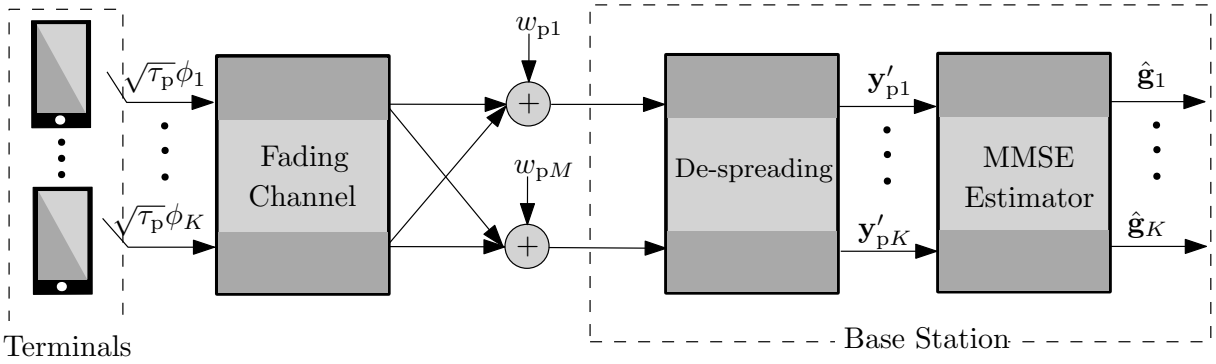



Figure 3.5: Channel Estimation via uplink pilot transmission.

### 3.2.3 Pilot contamination

Owing to the limited availability of frequency spectrum, there is a practical necessity to reuse the time-frequency resources across cells. In particular, when the resources designed for uplink pilot training are shared, the resulting channel estimate and thus system performance is affected by *pilot contamination*. Indeed it is not possible to assign orthogonal pilot sequences for all users in all cells, due to the limitation of the channel coherence interval. Orthogonal pilot sequences have to be reused from cell to cell. Therefore, the channel estimate obtained in a given cell will be contaminated by pilots transmitted by users in other cells [63]. Pilot contamination has an important impact beyond channel estimation, since the contamination makes it particularly hard for the BS to mitigate interference between terminals that use the same pilot. Pilot contamination is often described as a major inherent limitation of massive MIMO, as the effect persists even when the number of BS antennas grows to infinity [64]. Considerable efforts have been made to reduce this effect. The eigenvalue-decomposition-based channel estimation, pilot decontamination, as well as pilot contamination precoding schemes are proposed in [65–69].

## 3.3 Downlink Data Transmission

The BS wants to transmit the messages to each terminal, but it has to ensure that each terminal receives only the message intended to it. To do so, the BS can combine the channel estimates, obtained in the uplink training phase, with the vector comprising the messages designated to the terminals to obtain the actual signal to be transmitted. This combination is also known as precoding. Basically, when the message vector is obtained via a linear mapping, i.e., by multiplying $\mathbf{q}$ by a matrix we are performing a linear precoding. In contrast, when there is a nonlinear mapping, the employed precoder is nonlinear. We consider the linear precoding due to its low complexity. However, in chapter 4 we mention an algorithm that employs a nonlinear precoding scheme. Figure 3.6 illustrates the process of transmitting data in the downlink direction.

### 3.3.1 Linear Precoding

Linear precoding is a technique in which the signal is weighted at the transmitter so that interference among users is minimized in the downlink reception. Indeed, the precoding vectors determine the spatial directivity of the transmission. The BS

transmits the downlink signal

$$\mathbf{x} = \mathbf{P}\text{diag}\left(\boldsymbol{\eta}\right)^{1/2}\mathbf{q} = \sum_{k=1}^{K}\sqrt{\eta_k}\mathbf{p}_k q_k \tag{3.23}$$

in which $q_k$ is the message intended for the $k$th terminal. The transmit power allocated to terminal $k$ is denoted by $\eta_k$ so that

$$\sum_{k=1}^{K}\eta_k \leq 1. \tag{3.24}$$

In equation (3.23), $\mathbf{p}_k$ is the $k$-th column of the $M \times K$ precoding matrix $\mathbf{P}$, which is detailed in section 3.5. The precoding vectors $\mathbf{p}_k, k \in \{1, \cdots, K\}$, must satisfy $E\{\|\boldsymbol{p}_k\|^2\} = 1$, such that

$$E\{\sqrt{\eta_k}\left\|\boldsymbol{p}_k q_k\right\|^2\} = \eta_k. \tag{3.25}$$

Such a precoding normalization can be performed by making $\|\mathbf{p}_k\|^2 = 1, k \in \{1, \cdots, K\}$ [26].

The received signal $y_k$ at terminal $k$ is modeled as

$$\begin{aligned} y_k &= \mathbf{g}_k^{\mathrm{T}}\mathbf{x} + w_k \\ &= \sum_{k=1}^{K}\sqrt{\eta_k}\mathbf{g}_k^{\mathrm{T}}\mathbf{p}_k q_k + w_k \\ &= \underbrace{\mathbf{g}_k^{\mathrm{T}}\mathbf{p}_k\sqrt{\eta_k}q_k}_{\text{Desired signal}} + \underbrace{\mathbf{g}_k^{\mathrm{T}}\sum_{\substack{k'=1 \\ k'\neq k}}^{K}\mathbf{p}'_k\sqrt{\eta'_k}q'_k}_{\text{User-interference}} + \underbrace{w_k}_{\text{Noise}} \end{aligned} \tag{3.26}$$

in which $w_k$ is a realization of a random variable distributed as $\mathcal{CN}(0, \sigma_w^2)$ and represents the receiver noise.

## 3.4 Uplink Data Transmission

In this case, the BS wants to detect the signals transmitted from the $K$ terminals. When the number of BS antennas is large, linear decoders are usually good enough. With linear decoding schemes at the BS, the received signal $\mathbf{y}$ is separated into $K$ streams by multiplying it by a multiuser decoding matrix. Each stream is then decoded independently. The uplink data transmission procedure is depicted in Figure 3.7.

The $k$th terminal transmits a weighted symbol,

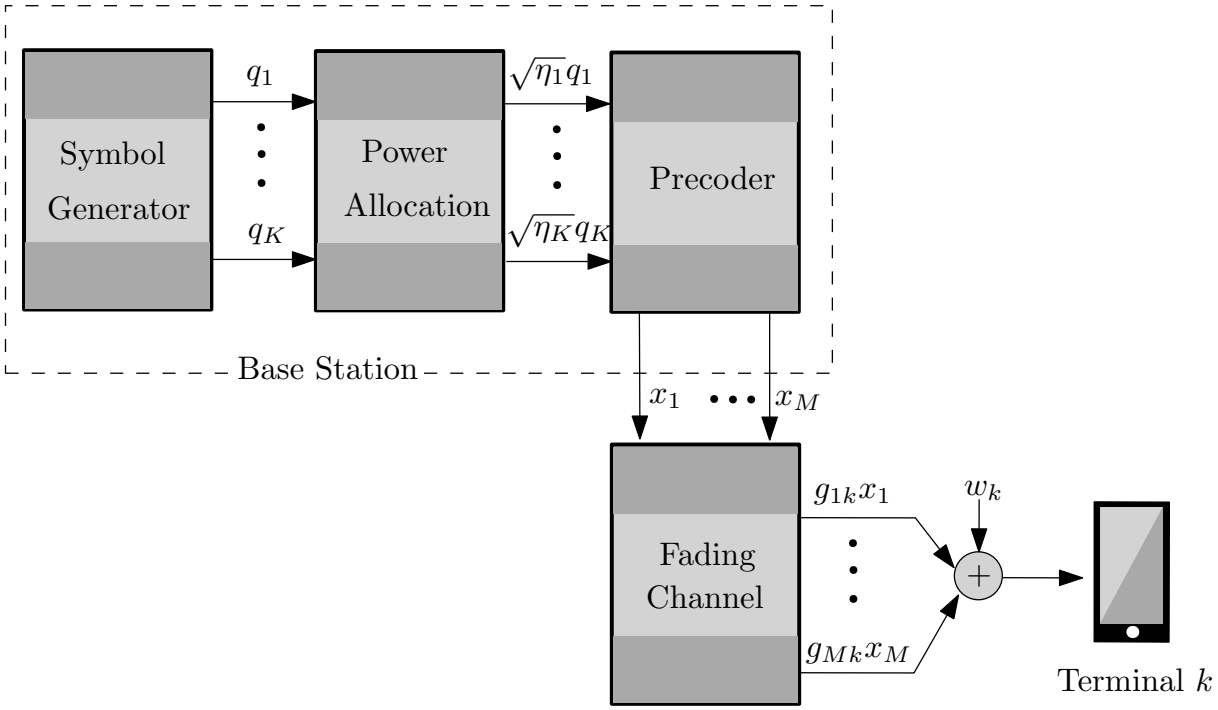



Figure 3.6: Downlink Data Transmission to terminal $k$.

$$x_k = \sqrt{\eta_k} q_k, \tag{3.27}$$

where $\eta_k$ is a power control coefficient. The symbols $q_k$ have zero mean and unit variance, and they are uncorrelated, so that

$$E[\boldsymbol{q}\boldsymbol{q}^{\mathrm{H}}] = \mathbf{I}_{(K)}, \tag{3.28}$$

where $\mathbf{q} = [q_1, \ldots, q_K]^T$ is the vector composed of the messages from all terminals. The signal received at the $m$th BS antenna element is a linear combination of the signals transmitted by all terminals

$$y_m = \sum_{k=1}^{K} g_{mk} x_k + w_m. \tag{3.29}$$

Hence, the complete $M \times 1$ received signal is

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}. \tag{3.30}$$

### 3.4.1 Linear Decoding

Linear decoding [70] is used to separate the signal $x_k$ of terminal $k$ from the other interfering terminals. To obtain an estimate of $x_k$, the BS pre-multiplies the

received signal $\mathbf{y}$ by the decoding vector $\mathbf{v}_k^{\mathrm{H}}$,

$$\hat{x_k} = \mathbf{v}_k^{\mathrm{H}}\mathbf{y} = \underbrace{\mathbf{v}_k^{\mathrm{H}}\mathbf{g}_k q_k}_{\text{Desired signal}} + \underbrace{\sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{v}_k^{\mathrm{H}}\mathbf{g}_i q_i}_{\text{User-interference}} + \underbrace{\mathbf{v}_k^{\mathrm{H}}\mathbf{w}}_{\text{Noise}}. \tag{3.31}$$

The decoding vector $\mathbf{v}_k$ is the $k$th column of the decoding matrix $\mathbf{V}$. The precoding matrix, as well as the decoding matrix are presented in section 3.5.

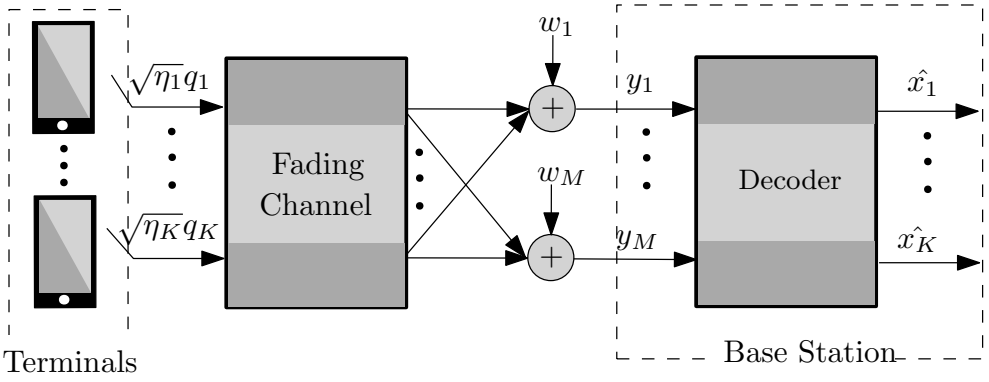



Figure 3.7: Uplink Data Transmission to BS.

## 3.5 Precoders/Decoders

One of the advantages of massive MIMO, mentioned in subsection 2.2.3, is the possibility of using simple signal processing due to the huge number of transmit antennas. Indeed, many linear processing approaches were provided in [71–75]. The most common linear precoding/decoding schemes are: maximum ratio and zero forcing.

### 3.5.1 Maximum Ratio

With the maximum ratio (MR), the BS wants to maximize the received SNR of each stream, ignoring the effect of multiuser interference [76, 77]. Then, the precoding matrix

$$\mathbf{P}_{\mathrm{MR}} = \mathbf{G}^* \tag{3.32}$$

is the conjugate channel matrix, whereas the decoding matrix

$$\mathbf{V}_{\mathrm{MR}} = \mathbf{G} \tag{3.33}$$

is the channel matrix. Since MR neglects the effect of multiuser interference, it performs poorly in interference-limited scenarios.

### 3.5.2 Zero Forcing

Zero forcing (ZF) is a simple method which decouples the multiuser channel into multiple independent subchannels and reduces the design to a power allocation problem. Indeed, ZF takes the interuser interference into account, but neglects the effect of noise. Thus, it performs very well in the high SNR regime, but it works poorly under noise-limited scenarios [78]. The precoding matrix,

$$\mathbf{P}_{\mathrm{ZF}} = \mathbf{G}^{*}(\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*})^{-1} \tag{3.34}$$

is the right pseudo-inverse of the channel matrix $\mathbf{G}^{\mathrm{T}}$, whereas the decoding matrix,

$$\mathbf{V}_{\mathrm{ZF}} = \mathbf{G}(\mathbf{G}^{\mathrm{H}}\mathbf{G})^{-1} \tag{3.35}$$

is the right pseudo-inverse of the channel matrix $\mathbf{G}$. Compared with MR, ZF has a higher implementation complexity due to the computation of the pseudo-inverse of the channel gain matrix.

## 3.6 Channel Hardening and Favorable Propagation

Channel hardening and favorable propagation are two related, but different properties. Channel hardening means that a fading channel behaves as a nearly deterministic channel [79]. Favorable propagation means that the channel vectors from different users are almost orthogonal [80]. These are both consequences of the law of large numbers.

### 3.6.1 Channel Hardening

Channel hardening makes a fading channel behave as deterministic, so that the random fluctuations in the channel due to microscopic changes in the propagation environment become negligible [81]. This property alleviates the need for combating small-scale fading and improves the downlink channel gain estimation. A propagation channel $\mathbf{g}_k$ provides asymptotic channel hardening if the gain $\|\mathbf{g}_s\|_2^2$ of the fading channel is close to its mean value as the number of antennas increases,

$$\frac{\|\boldsymbol{g}_k\|_2^2}{E[\|\boldsymbol{g}_k\|_2^2]} \to 1, M \to \infty. \tag{3.36}$$

Equation (3.36) also implies that

$$\mathrm{Var}\left[\frac{\|\boldsymbol{g}_k\|_2^2}{E[\|\boldsymbol{g}_k\|_2^2]}\right] = \frac{\mathrm{Var}\left[\|\boldsymbol{g}_k\|_2^2\right]}{\left(E\left[\|\boldsymbol{g}_k\|_2^2\right]\right)^2} \to 0, \quad M \to \infty. \tag{3.37}$$

Moreover, by applying lemma

$$E[(\boldsymbol{g}_k^{\mathrm{H}}\mathbf{I}\boldsymbol{g}_k)^2] = (\mathrm{tr}(\mathbf{I}\mathbf{R}_k))^2 + \mathrm{tr}(\mathbf{I}(\mathbf{R}_k)^2\mathbf{I}^{\mathrm{H}}), \tag{3.38}$$

from [26] in (3.36), we obtain

$$\begin{aligned}\frac{\mathrm{Var}\left[\|\boldsymbol{g}_k\|_2^2\right]}{\left(E\left[\|\boldsymbol{g}_k\|_2^2\right]\right)^2} &= \frac{(\mathrm{tr}(\mathbf{R}_k))^2 + \mathrm{tr}((\mathbf{R}_k)^2) - (\mathrm{tr}(\mathbf{R}_k))^2}{(\mathrm{tr}(\mathbf{R}_k))^2} \\ &= \frac{\mathrm{tr}((\mathbf{R}_k)^2)}{(\mathrm{tr}(\mathbf{R}_k))^2} = \frac{\mathrm{tr}((\mathbf{R}_k)^2)}{(M\beta_k)^2}.\end{aligned} \tag{3.39}$$

Therefore, if the variance in (3.39) is not close to zero as $M$ increases, one can assure that the channel does not harden [79]. Thus, we can compute the variance in equation (3.39) for both uncorrelated and correlated fading channels to check if channel hardening is observed. First, observe that $\mathrm{tr}((\mathbf{R}_k)^2)$ in the numerator of (3.39) is the sum of the squared eigenvalues of $\mathbf{R}_k$. In the case of uncorrelated fading, all the eigenvalues are equal to $\beta_k$ and hence $\mathbf{R}_k = \beta_k\mathbf{I}_{(M)}$, where $\beta_k$ is the coefficient regarding the large-scale fading effect. Thus, equation (3.39) becomes

$$\begin{aligned}\frac{\mathrm{tr}((\mathbf{R}_k)^2)}{(M\beta_k)^2} = \frac{\mathrm{tr}((\beta_k\mathbf{I}_{(M)})^2)}{(M\beta_k)^2} &= \frac{M\beta_k^2}{(M\beta_k)^2} \\ &= \frac{1}{M} \to 0, \quad M \to \infty\end{aligned} \tag{3.40}$$

which confirms the presence of channel hardening. On the other hand, in the case of correlated fading, $\mathbf{R}_k$ is not diagonal in general. Indeed, strong spatial correlation is characterized by large eigenvalues variations which thereby reduce the level of channel hardening that is observed for a given number of antennas. Then, more antennas are required to achieve a certain value in (3.39) under spatially correlated fading than with uncorrelated fading.

### 3.6.2 Favorable Propagation

Favorable propagation is observed if the channel vector $\mathbf{g}_k$, for $k = 1, \ldots K$ are pairwisely orthogonal [80], that is, if

$$\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{g}_{k'} \begin{cases} 0, \quad k,k' = 1, \ldots K, \quad k \neq k' \\ \|\boldsymbol{g}_k\|_2^2 \neq 0, \quad k = 1, \ldots K. \end{cases} \tag{3.41}$$

Unfortunately, the condition in (3.41) is not true in practice. However, we can investigate if the channel offers approximately favorable propagation. The pair of

channels $\mathbf{g}_k$ and $\mathbf{g}_{k'}$ provide asymptotically favorable propagation if

$$\frac{\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{g}_{k'}}{\sqrt{E[\|\boldsymbol{g}_k\|_2^2]E[\|\boldsymbol{g}_{k'}\|_2^2]}} \to 0, M \to \infty \tag{3.42}$$

This means that the inner product of the normalized channels $\boldsymbol{g}_{k'}/\sqrt{E[\|\boldsymbol{g}_k\|_2^2]}$ and $\boldsymbol{g}_k/\sqrt{E[\|\boldsymbol{g}_{k'}\|_2^2]}$ goes asymptotically to zero. For correlated Rayleigh fading channels, a sufficient condition for (3.42) is that the spatial correlation matrices $\mathbf{R}_k$ and $\mathbf{R}_{k'}$ have spectral norms that are bounded and the average channel gains remain strictly positive as $M \to \infty$. Notice that under this condition, the two channels will also exhibit asymptotic channel hardening. Likewise channel hardening, one can consider the variance

$$\mathrm{Var}\left[\frac{\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{g}_{k'}}{\sqrt{E[\|\boldsymbol{g}_k\|_2^2]E[\|\boldsymbol{g}_{k'}\|_2^2]}}\right] = \frac{\mathrm{tr}(\mathbf{R}_k\mathbf{R}_{k'})}{\mathrm{tr}(\mathbf{R}_k)\,\mathrm{tr}(\mathbf{R}_{k'})} = \frac{\mathrm{tr}(\mathbf{R}_k\mathbf{R}_{k'})}{M^2\beta_k\beta_{k'}} \to 0 \tag{3.43}$$

of the expression in (3.42) and verify if it holds. The closer to zero the variance is, the more orthogonal the channel directions are and less interference the terminals cause to each other. Ideally, the variance in (3.43) should be zero, but in practice it is not (unfavorable propagation). Therefore, using precoding/decoding schemes is still a good choice to mitigate interuser interference.

If both channels have uncorrelated fading, the variance becomes $1/M$ and thus decreases with an increasing number of antennas. In general, it is the spatial channel correlation that determines the variance in (3.43). It is zero if the terminals have orthogonal correlation eigenspaces, while the worst-case appears when the terminals have identical eigenspaces and only a few strong eigenvalues.

## 3.7 Concluding Remarks

In this chapter, we focused on the massive MIMO system. The system model was introduced, and the uplink and downlink modes were also mathematically described. Furthermore, the channel hardening and favorable propagation aspects of massive MIMO systems were briefly discussed. The next chapter exposes how to perform antenna selection in the downlink of a massive MIMO system.

# Chapter 4

# Antenna Selection in Massive MIMO

In theory, massive MIMO systems provide an impressive improvement in performance in comparison to single-input single-output (SISO) counterparts in terms of link reliability, data rate and radiated-energy efficiency due to the large number of BS-antennas. Unfortunately, increasing the number of antennas at BS may lead to undesired costs in practice. MIMO systems with $M$ BS-antennas require $M$ RF chains at the transmitter, which include low-noise amplifiers, downconverters, and digital-to-analog converters. Therefore, schemes in which the most suitable $S$ out of $M$ antennas are selected to be active have been proposed to reduce the number of required RF chains from $M$ to $S$ [20, 82, 83]. Antenna selection strategies lead to significant savings, but a performance loss is observed when compared to the full system. The main goal is to select the subset of antennas that preserves the system performance at a certain level.

In this chapter, we introduce the problem of selecting $S$ out of $M$ BS antennas in the downlink of a single-cell massive MIMO system as illustrated in Figure 4.1.

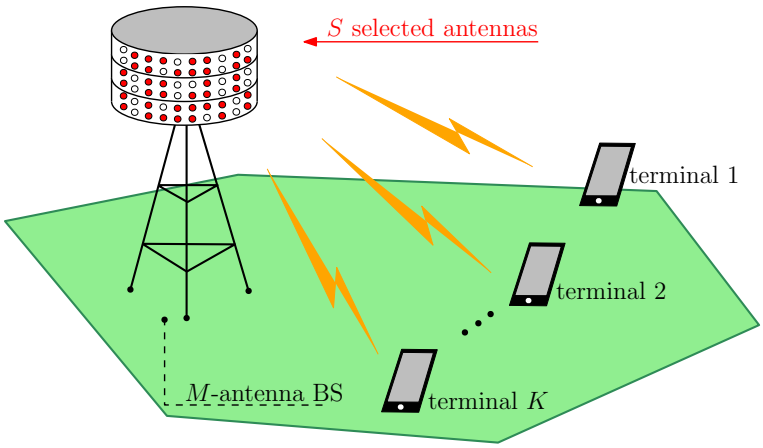



Figure 4.1: Downlink massive MIMO with only $S$ selected BS-antennas.

## 4.1 Antenna Selection

When all the BS antennas are active, the data transmission in the downlink of a massive MIMO system is modelled as in equation (3.3) and is also illustrated in Figure 3.6. To model the selection of a subset of $S$ antennas to transmit the data, we add a new step at the BS processing. This step, depicted in Figure 4.2, comprises the antenna selection algorithm which generates matrix $\mathbf{G}_S \in \mathbb{C}^{S\times K}$. The subscript $S$ in matrix $\mathbf{G}_S$ denotes that $S$ rows are selected out of $M$ in the full channel matrix $\mathbf{G}$. In general, the antenna selection algorithms only have the estimated channel matrix as input, so that the switch in Figure 4.2 is opened. This means that the antenna selection procedure is only needed when the channel matrix changes. However, some antenna selection strategies also require the so-called intended/desired vector $\mathbf{q}$ as input, hence the switch in Figure 4.2 is closed. In this case, the antenna selection procedure has to be performed whenever a new intended vector $\mathbf{q}$ is generated. Hence, the precoding matrix is recomputed every time a desired vector $\mathbf{q}$ is generated. Such a precoder is known as *symbol-level precoder* [84].

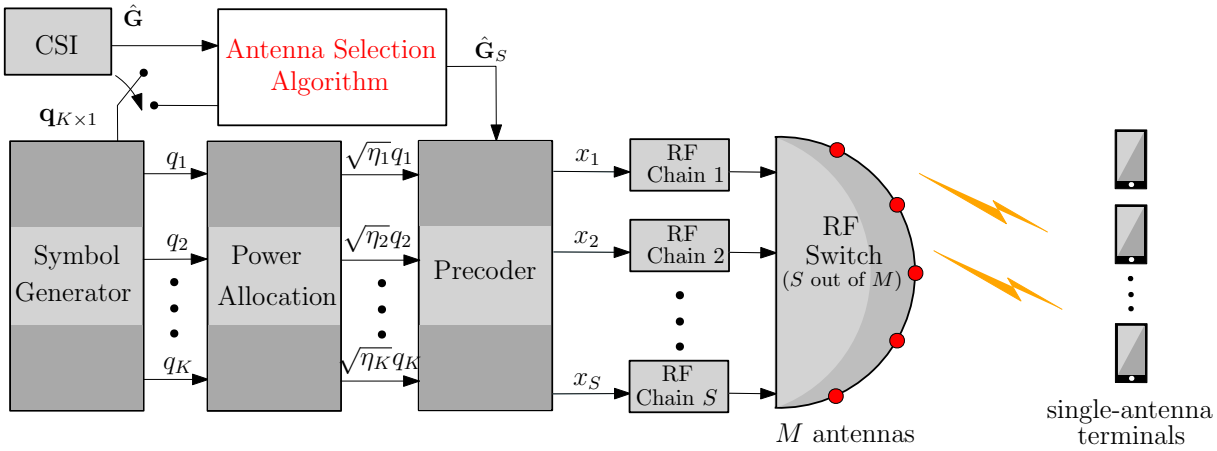



Figure 4.2: Antenna selection process in the downlink of a massive MIMO system.

### 4.1.1 Problem Description

Based on some selection criterion, the antenna selection algorithms produce the selection vector,

$$\mathbf{z} = [z_1,\ z_2,\ \ldots,\ z_M]^{\mathrm{T}} \in \{0,1\}^M \tag{4.1}$$

in which $z_m = 1$ denotes that the antenna with index $m$ was selected. Moreover, the selection vector must satisfy

$$\mathbf{1}^T\mathbf{z} = S, \tag{4.2}$$

where $S$ is the number of selected antennas. We can form the $S$-selected channel matrix,

$$\mathbf{G}_S = \text{rem}(\text{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\text{T}}\mathbf{G}, \tag{4.3}$$

where $\mathbf{Z}$ is the antenna selection matrix. We can obtain matrix $\mathbf{Z}$ by removing the zero columns of the diagonal matrix $\text{diag}\,(\mathbf{z})$. Indeed, the $\text{rem}(\cdot)$ operator represents the action of removing the zero columns of a given matrix, and is defined in Table 1.1. The message $\mathbf{x}_S \in \mathbb{C}^{S\times 1}$

$$\mathbf{x}_S = \mathbf{P}_S \text{diag}\,(\boldsymbol{\eta})^{1/2}\,\mathbf{q}, \tag{4.4}$$

is transmitted by the $S$ active BS antennas, where $\mathbf{P}_S = \mathbf{G}_S^*(\mathbf{G}_S^{\text{T}}\mathbf{G}_S^*)^{-1}$ with $S \geq K$ for ZF precoding and $\mathbf{P}_S = \mathbf{G}_S^*$ for MR precoding. The received signal in the $k$th terminal is of the form,

$$y_k = \mathbf{g}_{S_k}^{\text{T}}\mathbf{x}_S + w_k, \tag{4.5}$$

where $\mathbf{g}_{S_k}$ is the $k$th column of $\mathbf{G}_S$.

### 4.1.2 Selection Criteria

The most common criterion used to select the antennas is the one that aim at maximizing the downlink capacity [20, 85–87]. In addition, a recent method exploits the sparse recovery problem in the precoding stage, so that antenna selection and precoding procedures are jointly performed [21]. Our proposed methods follow a similar strategy, but at lower complexity.

#### 4.1.2.1 Capacity maximization

The maximization of the downlink capacity problem, detailed in equation (2.6), is rewritten to account for the antenna selection as

$$C_{\text{dl}} = \max_{\substack{\eta_k \geq 0 \\ \sum_{k=1}^{K}\eta_k \leq 1}} \log\det\left(\mathbf{I}_{(M)} + \rho_{\text{dl}}\mathbf{Z}^{\text{T}}\mathbf{G}\text{diag}\,(\boldsymbol{\eta})\,\mathbf{G}^{\text{H}}\mathbf{Z}\right), \tag{4.6}$$

where $\text{diag}\,(\boldsymbol{\eta})$ allocates the power among the $K$ terminal channels. By considering $\mathbf{A} = \mathbf{Z}^{\text{T}}\mathbf{G} \in \mathbb{C}^{S\times K}$ and $\mathbf{B} = \text{diag}\,(\boldsymbol{\eta})\,\mathbf{G}^{\text{H}}\mathbf{Z} \in \mathbb{C}^{K\times S}$ we can use the Sylvester's determinant theorem $\det(\mathbf{I}_{(S)} + \mathbf{AB}) = \det(\mathbf{I}_{(K)} + \mathbf{BA})$ to rewrite (4.6) as

$$C_{\text{dl}} = \max_{\substack{\eta_k \geq 0 \\ \sum_{k=1}^{K}\eta_k \leq 1}} \log\det\left(\mathbf{I}_{(K)} + \rho_{\text{dl}}\text{diag}\,(\boldsymbol{\eta})\,\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G}\right), \tag{4.7}$$

where $\mathbf{Z}\mathbf{Z}^{\mathrm{T}} = \mathrm{diag}\,(\mathbf{z})$. Then, we can formulate the optimization problem,

$$\begin{aligned} \underset{\substack{\mathbf{z}\in\mathbb{R}^{M\times 1}\\ \boldsymbol{\eta}\in\mathbb{R}^{K\times 1}}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \rho_{\mathrm{dl}}\mathrm{diag}\,(\boldsymbol{\eta})\,\mathbf{G}^{\mathrm{H}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G})) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad z_m \in \{0,1\}, m \in \{1,\cdots,M\} \\ & \mathbf{1}^{\mathrm{T}}\boldsymbol{\eta} = 1 \quad \eta_k \geq 0, k \in \{1,\cdots,K\} \end{aligned} \tag{4.8}$$

in which $\boldsymbol{\eta}$ is the power allocation vector.

#### 4.1.2.2 Sparse Recovery

The antenna selection problem can also be interpreted as finding the most significant columns of a dictionary matrix $\mathbf{D} \in \mathbb{C}^{M\times K}$ that is, the columns that most contribute to the recovering of the desired vector $\mathbf{b} \in \mathbb{C}^{K\times 1}$. Such a problem is known as *sparse recovery* and can be formulated as

$$\begin{aligned} \underset{\mathbf{a}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{D}\mathbf{a} - \mathbf{b}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{a}\|_0 = S \end{aligned} \tag{4.9}$$

where $S$ is the number of non-zero elements of vector $\mathbf{a}$. We can choose the dictionary matrix equal to the channel matrix $\mathbf{D} = \mathbf{G}$, the desired vector equals the intended signal $\mathbf{b} = \mathbf{q}$, and the sparse vector equals the transmitted message $\mathbf{a} = \mathbf{x}$. The equivalent optimization problem is

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{G}\mathbf{x} - \mathbf{q}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{x}\|_0 = S \end{aligned} \tag{4.10}$$

in which $S$ represents the number of selected antennas. Essentially, the problem in (4.10) consists of finding a sparse precoding scheme so that a good estimate of vector $\mathbf{q}$ is obtained at the terminal.

### 4.1.3 Convex Optimization Methods for Antenna Selection

In the massive MIMO system setup, the subset of active antennas is chosen to be either the one that maximizes the downlink capacity or the one that minimizes the bit error rate at reception. Unfortunately, those problems are not convex due to the presence of binary and $l_0$-norm constraints, which are non-convex. A near-optimal solution can be achieved thanks to convex relaxation and $l_1$-norm approximation. In this subsection, the antenna selection problems are rewritten using these techniques so that a convex optimization algorithm can be directly employed.

#### 4.1.3.1 Massive MIMO Downlink Capacity Convex Problem

The problem represented in equation (4.8) is over two vector variables, $\mathbf{z}$ and $\boldsymbol{\eta}$. In case that the entries of these vectors are continuous, the objective function in equation (4.8) is concave. However, the elements of $\mathbf{z}$ are binary integer variables and hence, the problem in equation (4.8) is not convex [20, 88]. We can replace the binary constraints of the form $z_m \in \{0,1\}$ in problem (4.8) with the convex constraints $z_m \in [0,1]$, and then obtain the following convex relaxed optimization problem [20, 88, 93, 94]

$$\begin{aligned} \underset{\substack{\mathbf{z}\in\mathbb{R}^{M\times 1}\\ \boldsymbol{\eta}\in\mathbb{R}^{K\times 1}}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \rho_{\text{dl}}\text{diag}\left(\boldsymbol{\eta}\right)\mathbf{G}^{\text{H}}\text{diag}\left(\mathbf{z}\right)\mathbf{G})) \\ \text{subject to} \quad & \mathbf{1}^{\text{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, m \in \{1, \cdots, M\} \\ & \mathbf{1}^{\text{T}}\boldsymbol{\eta} = 1 \quad \eta_k \geq 0, k \in \{1, \cdots, K\} \end{aligned} \tag{4.11}$$

over two vector variables. In order to simplify the two sets of variables problem in equation (4.11), the optimization is divided into two steps [20]. In the first step, we assume equal power allocation among the users so that $\eta_k = 1/K$ for $k \in \{1, \cdots, K\}$. Then resulting the optimization problem

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\text{H}}\text{diag}\left(\mathbf{z}\right)\mathbf{G})) \\ \text{subject to} \quad & \mathbf{1}^{\text{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, m = 1, \ldots M \end{aligned} \tag{4.12}$$

is over the selection vector. The optimization problem in equation (4.12) can be solved by using interior-point methods [50]. Alternatively, we can obtain an approximate version of the problem by using one iteration of the log-barrier method [88]. Using this strategy, the inequality constraints are implicit in the objective function of the problem,

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\text{H}}\text{diag}\left(\mathbf{z}\right)\mathbf{G})) + \\ & \kappa\sum_{m=1}^{M}(\log(z_m) + \log(1 - z_m)) \\ \text{subject to} \quad & \mathbf{1}^{\text{T}}\mathbf{z} = S \end{aligned} \tag{4.13}$$

where $\kappa$ is a positive constant which controls the quality of the approximation. It is worth mentioning that the log-barrier method is also a interior-point method based on successive approximations. However, here we consider only one approximation and hence the value of $\kappa$ is fixed. The optimization problem in (4.13) can be solved via Newton Method [50] as we only have equality constraints. Instead of a binary vector, the solution of problems (4.11), (4.12) and (4.13) is a vector whose elements

are between zero and one. In order to obtain a binary vector, which suits better a vector to represent the antenna selection, the $S$ largest elements of $\mathbf{z}$ can be replaced by one and the remaining by zero. The resulting $\mathbf{z}$ is used to compute $\mathbf{G}_S$, the channel matrix after selection, as in equation (4.3). The antenna selection algorithm based on the problem in equation (4.12) is presented in Algorithm 1. The sets $\mathcal{I}$ and $\mathcal{I}'$ in Algorithm 1 are meant to store the indices of the selected and non-selected antennas, respectively.

In the second step, we optimize over the user power allocation and thus find the maximum average capacity

$$\begin{aligned} &\underset{\boldsymbol{\eta}\in\mathbb{R}^{K\times 1}}{\text{maximize}} && \log_2(\det(\mathbf{I}_{(K)} + \rho_{\text{dl}}\text{diag}\left(\boldsymbol{\eta}\right)\mathbf{G}_S{}^{\text{H}}\mathbf{G}_S)) \\ &\text{subject to} && \mathbf{1}^T\boldsymbol{\eta} = 1 \quad \eta_k \geq 0, k = 1, \dots K \end{aligned} \tag{4.14}$$

by using water-filling algorithms [50, 51].

**Algorithm 1** : Max Capacity Antenna Selection (MCAS) [20]

1) Input: $\mathbf{G}$, $S$
2) Initialization: $\mathcal{I}' = \{1, 2, \dots M\}$, $\mathcal{I} = \{\}$
3) Compute $\mathbf{z}$ that solves the problem (equation (4.12))

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} && \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\text{H}}\text{diag}\left(\mathbf{z}\right)\mathbf{G})) \\ &\text{subject to} && \mathbf{1}^{\text{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, m = 1, \dots M \end{aligned}$$

by employing interior-point methods [50].
4) Store the indices of the $k$-largest $\mathbf{z}$ in $\mathcal{I}$
5) $\mathcal{I}' = \mathcal{I}' - \mathcal{I}$
6) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
7) Set $z_i \leftarrow 0$ for $i \in \mathcal{I}'$
8) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \text{rem}(\text{diag}\left(\mathbf{z}\right))\mathbf{G} = \mathbf{Z}^{\text{T}}\mathbf{G}$$

9) Output: $\mathbf{G}_S$

#### 4.1.3.2 Massive MIMO LASSO Convex Problem

Although the $l_0$-norm leads to a sparse solution, it produces a challenging optimization problem to solve, due to its discontinuity and non-convexity. Hence, even though the objective function of the problem in equation (4.10) is convex, the $l_0$-norm constraint produces a non-convex optimization problem. The $l_1$-norm, a continuous and convex surrogate, can be used as an approximation of the $l_0$-norm [95]. By replacing the $l_0$-norm with the $l_1$-norm and the equality with inequality in

equation (4.10), we obtain

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{G}\mathbf{x}-\mathbf{q}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{x}\|_1 \leq \gamma_1 \end{aligned} \tag{4.15}$$

which is known as least absolute shrinkage and selection operator (LASSO) method [96]. The constraint $\|\mathbf{x}\|_1 \leq \gamma_1$ induces a sparse precoding vector $\mathbf{x}$. To ensure that the power used to transmitting $\mathbf{x}$ does not exceed the transmit power used when all the antennas are active, a new constraint is added. By adding such a constraint we end up with the antenna selection problem stated in [21]

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{G}\mathbf{x}-\mathbf{q}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{x}\|_1 \leq \gamma_1 \qquad\qquad \|\mathbf{x}\|_2 \leq \gamma_2 \\ & \text{where} \;\; \gamma_1 = \alpha_1 \|\mathbf{x}_{\text{ZF}}\|_1 , \;\; \gamma_2 = \alpha_2 \|\mathbf{x}_{\text{ZF}}\|_2 . \end{aligned} \tag{4.16}$$

The parameters $\gamma_1$ and $\gamma_2$ are chosen as fraction of the $l_1$ and $l_2$ norms of the zero forcing solution

$$\mathbf{x}_{\text{ZF}} = \mathbf{P}_{\text{ZF}}\mathbf{q} \tag{4.17}$$

where $\mathbf{P}_{\text{ZF}}$ is the ZF precoding matrix defined in equation (3.34). By setting parameter $\alpha_2 = 1$, the $l_2$-norm constraint works like a energy controller of the precoded signal $\mathbf{x}$. Parameter $\alpha_1 \in [0,1]$ is called *sparsity factor* and is used to control the sparsity of the solution. The value of the parameter $\alpha_1$ does not correspond to the percentage of active antennas, but it is closely related as reported in [21]. Problem in equation (4.16) gives rise to Algorithm 2 which is the joint-precoding and antenna selection algorithm proposed in [21]. Hence, the algorithm's output is the precoding vector $\mathbf{x}$ itself rather than the $S$-selected channel matrix. Indeed, the algorithm in Algorithm 2 induces a nonlinear mapping as there is no precoding matrix $\mathbf{P}$ that linearly delivers precoding vector $\mathbf{x}$. Moreover, such an algorithm is a *symbol-level precoder*, which means that its solution varies with the symbol vector $\mathbf{q}$.

**Algorithm 2** : LASSO [21]

1) Input: $\mathbf{G}$, $\alpha_1$, $\mathbf{q}$
2) Compute the ZF solution, $\mathbf{x}_{\text{ZF}} = \mathbf{G}^*(\mathbf{G}^{\text{T}}\mathbf{G}^*)^{-1}\mathbf{q}$
3) Find $\mathbf{x}$ that solves the problem (equation (4.16))

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{G}\mathbf{x}-\mathbf{q}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{x}\|_1 \leq \gamma_1 \qquad\qquad \|\mathbf{x}\|_2 \leq \gamma_2 \\ & \text{where} \;\; \gamma_1 = \alpha_1 \|\mathbf{x}_{\text{ZF}}\|_1 , \;\; \gamma_2 = \alpha_2 \|\mathbf{x}_{\text{ZF}}\|_2 . \end{aligned}$$

5) Output: $\mathbf{x}_{\text{LASSO}} \leftarrow \mathbf{x}$

### 4.1.4 Random Antenna Selection Problem

In the Algorithm 3, we consider that the active antennas are chosen randomly. Although it is the antenna selection algorithm with the least computational cost, it is also the antenna selection with lowest accuracy. Then, it is only used in this work to provide a performance lower bound.

---

**Algorithm 3** : Random Antenna Selection (RAS)

---

1) Input: $\mathbf{G}$
2) Initialization: $\mathbf{z} = \mathbf{0}_M$, $\mathcal{I} = \{\}, \mathcal{I}' = \{1,2,\dots M\}$
3) Choose randomly $S$ indices of $\mathcal{I}'$ and use them to form $\mathcal{I}$
5) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
6) Compute the $S$-selected channel matrix
$\mathbf{G}_S = \text{rem}(\text{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\text{T}}\mathbf{G}$
7) Output: $\mathbf{G}_S$

---

## 4.2 Concluding Remarks

In this chapter, the problem of selecting a subset of BS antennas was addressed. As a result, basic optimization problems concerning the antenna selection were stated. We briefly summarized the existing antenna selection algorithms based on convex optimization and also an algorithm based on random selection. In the next chapter, four proposed antenna selection algorithms are proposed based on the Matching Pursuit technique are introduced.

# Chapter 5

# Matching Pursuit for Antenna Selection

Overcoming the massive MIMO systems drawbacks regarding high computational complexity and BS cost is one of the major concerns in recent research. The antenna selection strategy has been considered an efficient approach to reduce the number of RF elements with negligible loss in performance. Nonetheless, the complexity of these antenna selection algorithms should also be considered, in order to alleviate the BS computational effort and hence the overall system cost. For this purpose, in this chapter, we propose two low complexity antenna selection algorithms derived from the matching pursuit technique, which are the main contributions of this work. The proposed algorithms are addressed for channel-level and symbol-level precoders, respectively.

## 5.1 Matching Pursuit

Often used in compression, denoising, and pattern recognition, matching pursuit (MP) is a greedy algorithm employed to represent a signal using a redundant dictionary [22, 23, 99]. Iteratively, the MP aims to solve the $l_0$-norm constrained problem in (4.9), which is conveniently repeated below,

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{D}\mathbf{z}-\mathbf{b}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S \end{aligned} \tag{5.1}$$

by approximating the target vector $\mathbf{b}$ using the dictionary matrix $\mathbf{D}$ and a sparse vector $\mathbf{z}$.

First, consider a vector $\mathbf{z}$ which is multiplied by matrix $\mathbf{D}$ to obtain the target vector $\mathbf{b}$ in a linear system, as depicted in Figure 5.1. The ultimate objective of the MP algorithm is to identify which elements of vector $\mathbf{z}$ are less informative and

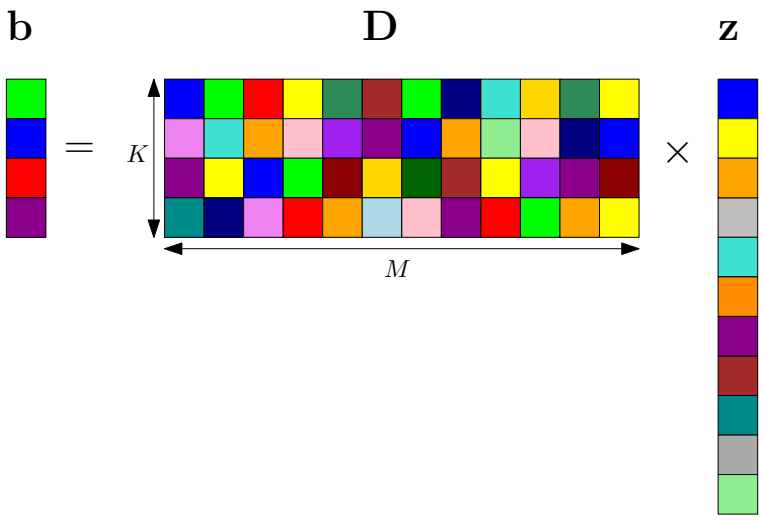



Figure 5.1: Example of linear system.

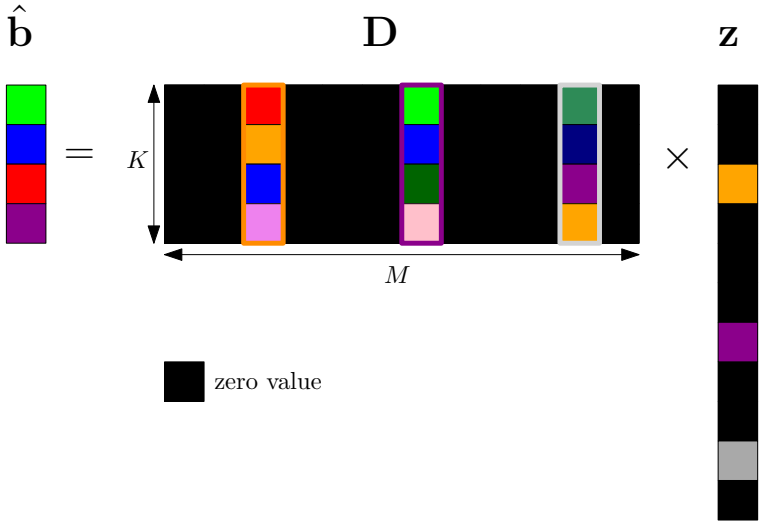



Figure 5.2: Example of *sparse recovery* problem.

hence can be set to zero, as it is outlined in Figure 5.2. Moreover, we can observe that zeroing the $i$th element of the vector $\mathbf{z}$ is equivalent to replace the $i$th column of the matrix $\mathbf{D}$ by a zero column, or selecting the most informative columns of matrix $\mathbf{D}$. The columns of $\mathbf{D}$, called codewords $\mathbf{c}_m, m \in \{1, \cdots M\}$, are vectors in the Hilbert space $\mathcal{H}$ [99]. Since we need to define distance, angles and orthogonality between signal vectors, we see the convenience of working in the Hilbert space. The Hilbert space is a Banach space where inner product exists [22]. A Banach space is a complete vector space with a metric that allows the computation of vector length and distance between vectors. As $\mathcal{H}$ is also a Banach space, it also admits the norm and must be complete. The completeness property entails the convergence of every *Cauchy sequence* in $\mathcal{H}$ to an element of $\mathcal{H}$. A sequence $f_n$ is a *Cauchy sequence* if for any $\epsilon > 0$ one has

$$\|f_n - f\| < \epsilon \tag{5.2}$$

when $n$ and $p$ are large enough [22].

Instead of approximating the $l_0$-norm constrained problem and applying a sophisticated optimization method or starting an exhaustive search over all possible combinations, the MP algorithm tries to find the best solution at each iteration. As a matter of fact, it searches for the codeword which is closest to the current residue. In the first iteration, the residue is the vector we want to approximate, whereas in the remaining iterations it is composed of the last residue and a modified version of the selected codeword. To measure how close the residue is from the codewords we use the inner product, represented by $\langle \cdot, \cdot \rangle$, so that the selected codeword with index $m_i$ achieves the maximum inner-product. The approximate version of $\mathbf{b}$ at the $i$th iteration

$$\widehat{\mathbf{b}}_i = \sum_{i=1}^{I} \langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle \mathbf{c}_{m_i} \tag{5.3}$$

is composed by the sum of the projections of the current residue in the selected codeword. Variable $I$ is the total number of iterations when the stopping criterion is reached. One of the possible stopping criteria is when the residue' norm is very close to zero. The inner product in equation (5.3) acts like a weight applied to each selected codeword, illustrated by the different colors in the selected column's border of the dictionary matrix $\mathbf{D}$ represented in Figure 5.2. The original corresponding algorithm is detailed in Algorithm 4.

**Algorithm 4** : Original Matching Pursuit (MP) [99]

1) Initialization
$i = 1$, $\mathbf{r}_1 = \mathbf{b}$, $\widehat{\mathbf{b}}_1 = \mathbf{0}_K$
2) Generate Dictionary
3) Repeat until a stop criterion is met:

$$\langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle = \max_{i \in \mathcal{I}} \{ |\langle \mathbf{r}_i, \mathbf{c}_m \rangle| \}$$

b) Choose $z_{m_i} = \langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle$
c) Let $\mathbf{r}_{i+1} = \mathbf{r}_i - z_{m_i} \mathbf{c}_{m_i}$
d) Let $\widehat{\mathbf{b}}_{i+1} = \widehat{\mathbf{b}}_i + z_{m_i} \mathbf{c}_{m_i}$
e) Increment $i$.
4)Stop
5)Output $\widehat{\mathbf{b}}_i$

## 5.2 Matching Pursuit as an Antenna Selection Strategy

Originally, the MP method aims at finding the sparsest vector $\mathbf{z}$ and thus the codewords can be chosen more than once, which is quite reasonable since the MP was conceived for compression applications. However, in the antenna selection problem, we are interested in selecting a fixed number of antennas, $S$. Then it is more convenient to employ an approach similar to the one used in the Orthogonal Matching Pursuit (OMP) technique [22], where a codeword can be chosen only once. Furthermore, by using the number of selected antennas $S$ as maximum number of iterations, we ensure that the $l_0$-norm of vector $\mathbf{z}$ is $S$.

The goal of the antenna selection problem is to obtain a subset of indices that represent the selected antennas. In fact, we are interested in the indices of the non-zero entries of $\mathbf{z}$, which are equivalent to the indices of the chosen codewords. We can then formulate the optimization problem that describes the antenna selection as

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{C}^{M\times 1}}{\text{minimize}} \quad & \|\mathbf{D}\mathbf{z}-\mathbf{b}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S, \quad \mathbf{z}\in\{0,1\}^M \end{aligned} \tag{5.4}$$

and solve it with a proposed version of MP, whose algorithm is described in Algorithm 5.

The main difference between the original MP algorithm and the proposed Matching Pursuit Antenna Selection (MPAS) algorithm is that $z_{m_i}$ is no longer the inner product value. Due to the additional constraint in problem (5.4), $\mathbf{z}\in\{0,1\}^M$, it is more convenient to have $z_{m_i}=1$. In problem (5.4), the stop criterion is the number $S$ of antennas we want to be active/selected. Thus the total number of iterations $I$ is equal to $S$. Another important issue to be mentioned is that selected codewords are no longer used at the remaining iterations. If the selected codewords are not discarded, we cannot guarantee that $S$ distinct antennas are selected.

The challenge now is to find which variables of the antenna selection problem can act like target vector $\mathbf{b}$ and dictionary matrix $\mathbf{D}$. We start by observing the signal received at a certain terminal. When all the BS antennas are active, the received signal at terminal $k$ is

$$\begin{aligned} y_k &= \mathbf{g}_k^{\mathrm{T}}\mathbf{x} + w_k \\ &= \mathbf{g}_k^{\mathrm{T}}\mathbf{P}\mathrm{diag}\left(\boldsymbol{\eta}\right)^{1/2}\mathbf{q} + w_k, \end{aligned} \tag{5.5}$$

where $\mathbf{P}$ is the precoding matrix and vector $\boldsymbol{\eta}$ contain the terminal power allocation.

**Algorithm 5** : Proposed Matching Pursuit Antenna Selection (MPAS)

1) Input: $\mathbf{b}$, $\mathbf{D} = [\mathbf{c}_1 \ \dots \ \mathbf{c}_M]$
2) Initialization: $i = 1$, $\mathbf{r}_1 = \mathbf{b}$, $\mathbf{z} = \mathbf{0}_M$, $\mathcal{I} = \{1, \dots M\}$
3) Repeat until $i = S - 1$:
a) Find the closest codeword, i.e., find $m_i \in \mathcal{I}$ such that

$$\langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle = \max_{i \in \mathcal{I}} \{ |\langle \mathbf{r}_i, \mathbf{c}_m \rangle| \}$$

b) Choose $z_{m_i} = 1$
c) Let $\mathbf{r}_{i+1} = \mathbf{r}_i - z_{m_i} \mathbf{c}_{m_i}$
d) $\mathcal{I} = \mathcal{I} - m_i$
e) Increment $i$.
4)Stop.
5) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \text{rem}(\text{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\text{T}}\mathbf{G}$$

6) Output: $\mathbf{G}_S$

In contrast, when only $S$ antennas are active the terminal $k$ receives

$$\begin{aligned} y'_k &= \mathbf{g}_{Sk}^{\text{T}} \mathbf{x}_S + w_k \\ &= \mathbf{g}_k^{\text{T}} \mathbf{Z} \mathbf{P}_S \text{diag}\,(\boldsymbol{\eta})^{1/2} \mathbf{q} + w_k, \end{aligned} \tag{5.6}$$

where $\mathbf{P}_S$ is the precoding matrix when $S$ antennas are active. Matrix $\mathbf{Z} \in \mathbb{C}^{M \times S}$ is obtained by removing the zero columns of matrix $\text{diag}\,(\mathbf{z})$, previously defined in subsection 4.1.1. By comparing equations (5.5) and (5.6), we see that if $\mathbf{Z}\mathbf{P}_S \approx \mathbf{P}$, then $y_k \approx y'_k$, which means that the obtained vector $\mathbf{z}$ leads to the minimum reception absolute error, $\|y_k - y'_k\|_2$. Thus our main goal is finding $\mathbf{z}$ with $\|\mathbf{z}\|_0 = S$ that leads to

$$\mathbf{Z}\mathbf{P}_S \approx \mathbf{P} \tag{5.7}$$

for the MR and ZF precoding schemes which were introduced in section 3.5.

### 5.2.1 MR Matching Pursuit Antenna Selection

In this subsection we propose an antenna selection algorithm based on matching pursuit, assuming that an MR precoding is performed. The matrix $\mathbf{P}$, previously defined in equation (3.32) as

$$\mathbf{P} = \mathbf{G}^*, \tag{5.8}$$

is the MR precoding matrix when all the BS antennas are active. In contrast, when only a subset of $S$ transmit antennas are active, the resulting precoding matrix

becomes

$$\begin{aligned} \mathbf{P}_S &= \mathbf{G}_S^* \\ &= \mathbf{Z}^{\mathrm{T}}\mathbf{G}^* \end{aligned} \tag{5.9}$$

By replacing $\mathbf{P}$ and $\mathbf{P}_S$ in equation (5.7), it follows that

$$\begin{aligned} \mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{G}^* &\approx \mathbf{G}^* \\ \mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^* &\approx \mathbf{G}^* \end{aligned} \tag{5.10}$$

where $\mathbf{Z}\mathbf{Z}^{\mathrm{T}} = \mathrm{diag}\,(\mathbf{z})$. This means that finding the vector $\mathbf{z}$ that makes $\mathbf{Z}\mathbf{P}_S \approx \mathbf{P}$ is equivalent to finding the one that ensures

$$\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^* \approx \mathbf{G}^{\mathrm{T}}\mathbf{G}^* \tag{5.11}$$

In fact we want to iteratively find the $z_m's$ that contribute more to

$$\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^* = \sum_{m=1}^{M} z_m(\mathbf{g}_m^{\mathrm{T}})^{\mathrm{T}}(\mathbf{g}_m^{\mathrm{T}})^* = \sum_{m=1}^{M} z_m\mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}, \tag{5.12}$$

where $\mathbf{g}_m^{\mathrm{T}}$ is the $m$th row of channel matrix $\mathbf{G}$, defined in equation (3.1). For that, it is convenient to express the approximation in equation (5.11) as

$$\begin{bmatrix} \mathbf{g}_1\mathbf{g}_1^{\mathrm{H}} & \mathbf{g}_2\mathbf{g}_2^{\mathrm{H}} & \dots \mathbf{g}_M\mathbf{g}_M^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} z_1\mathbf{I}_K \\ z_2\mathbf{I}_K \\ \vdots \\ z_m\mathbf{I}_K \end{bmatrix} \approx \sum_{m=1}^{M} \mathbf{g}_m\mathbf{g}_m^{\mathrm{H}} \tag{5.13}$$

due to the MP structure problem. The MP technique considers the approximation of a vector, however, in equation (5.13) we are trying to obtain an approximation for a matrix. Therefore, some manipulations are necessary so that the MP technique can be applied. We start by considering $\mathbf{A} = \sum_{m=1}^{M} \mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}$ and $\mathbf{D}_{\mathrm{aux}}$ as the following auxiliary block matrix

$$\mathbf{D}_{\mathrm{aux}} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \dots \mathbf{C}_M \end{bmatrix} \tag{5.14}$$

in which $\mathbf{C}_m = \mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}$, for $m \in \{1, \cdots, M\}$. Also, let $\mathbf{R}_1 = \mathbf{A}$ so that we can try to find the best match for $\mathbf{R}_1$ in $\mathbf{D}_{\mathrm{aux}}$ by computing

$$|\langle \mathbf{R}_1, \mathbf{C}_m \rangle| = \mathrm{tr}(\mathbf{R}_1^{\mathrm{H}}\mathbf{C}_m) \tag{5.15}$$

for $m \in \{1, \cdots, M\}$. Fortunately, the matrix inner product is equivalent to the vector inner product

$$\mathrm{tr}(\mathbf{R}_1^{\mathrm{H}}\mathbf{C}_m) = \mathrm{vec}(\mathbf{R}_1)^{\mathrm{H}}\,\mathrm{vec}(\mathbf{C}_m) = \mathbf{r}_1^{\mathrm{H}}\mathbf{c}_m. \tag{5.16}$$

The vec$(\cdot)$ operator takes a matrix as input, and outputs a column vector which is obtained by transposing the rows of the input matrix and stacking them up. We then propose the Algorithm 6 which is the Algorithm 5 with input vector $\mathbf{b} = \mathrm{vec}(\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*})$ and input matrix

$$\mathbf{D}_{\mathrm{MR}} = \left[\mathrm{vec}(\mathbf{g}_1\mathbf{g}_1^{\mathrm{H}}) \quad \mathrm{vec}(\mathbf{g}_2\mathbf{g}_2^{\mathrm{H}}) \quad \ldots \mathrm{vec}(\mathbf{g}_M\mathbf{g}_M^{\mathrm{H}})\right], \tag{5.17}$$

in which $\mathbf{C}_m = \mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}$. The corresponding sparse recovery for the antenna selection problem with MR precoding can then be expressed as

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} \quad & \left\|\mathbf{D}_{\mathrm{MR}}\mathbf{z} - \mathrm{vec}(\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*})\right\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S, \quad \mathbf{z} \in \{0,1\}^M \end{aligned} \tag{5.18}$$

and solved with Algorithm 6.

**Algorithm 6** : Proposed MR Matching Pursuit Antenna Selection (MR-MPAS)

1) Input: $\mathbf{b} = \mathrm{vec}(\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*})$, $\mathbf{D}_{\mathrm{MR}} = [\mathbf{c}_1 \ \ldots \ \mathbf{c}_M]$, $\mathcal{I} = \{1, \ldots M\}$
2) Initialization: $i = 1$, $\mathbf{r}_1 = \mathbf{b}$, $\mathbf{z} = \mathbf{0}_M$
3) Repeat until $i = S - 1$:
a) Find the closest codeword, i.e., find $m_i \in \mathcal{I}$ such that

$$\langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle = \max_{i\in\mathcal{I}} \{|\langle \mathbf{r}_i, \mathbf{c}_m \rangle|\}$$

b) Choose $z_{m_i} = 1$
c) Let $\mathbf{r}_{i+1} = \mathbf{r}_i - z_{m_i}\mathbf{c}_{m_i}$
d) $\mathcal{I} = \mathcal{I} - m_i$
e) Increment $i$.
4)Stop.
5) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

6) Output: $\mathbf{G}_S$

### 5.2.2 ZF Matching Pursuit Antenna Selection

In this subsection we propose an antenna selection algorithm based on matching pursuit, assuming that a ZF precoding is performed. The procedure prescribed in

the previous subsection 5.2.1 is adopted here for the ZF precoding scheme. The matrix, defined in equation (3.34) as

$$\mathbf{P} = \mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1}, \tag{5.19}$$

is the ZF precoding matrix when all the BS antennas are active. On the other hand, if only a subset of $S$ transmit antennas are active, the resulting precoding matrix becomes

$$\begin{aligned} \mathbf{P}_S &= \mathbf{G}_S^*(\mathbf{G}_S^{\mathrm{T}}\mathbf{G}_S^*)^{-1} \\ &= \mathbf{Z}^{\mathrm{T}}\mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathbf{Z}\mathbf{Z}^T\mathbf{G}^*)^{-1} \\ &= \mathbf{Z}^{\mathrm{T}}\mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1}. \end{aligned} \tag{5.20}$$

By replacing $\mathbf{P}$ and $\mathbf{P}_S$ in equation (5.7), we obtain

$$\begin{aligned} &\mathbf{Z}\mathbf{Z}^{\mathrm{T}}\mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1} \approx \mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1} \\ &\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1} \approx \mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1} \end{aligned} \tag{5.21}$$

which means that finding the vector $\mathbf{z}$ that makes $\mathbf{Z}\mathbf{P}_S \approx \mathbf{P}$ is equivalent to finding the one that ensures

$$\begin{cases} \mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^* \approx \mathbf{G}^* \\ (\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1} \approx (\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1}. \end{cases} \tag{5.22}$$

Observe that we can left-multiply the first statement in equation (5.22) by $\mathbf{G}^{\mathrm{T}}$ and obtain

$$\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^* \approx \mathbf{G}^{\mathrm{T}}\mathbf{G}^*, \tag{5.23}$$

which is equal to the MR objective statement in equation (5.11). Thus, the first statement is already the objective approximation for the MR precoding. For the sake of simplicity, we will only take the second statement in equation (5.22) into account. In fact, the statement

$$(\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1} \approx (\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1}. \tag{5.24}$$

can also be obtained if we impose the approximation

$$\mathbf{P}_S^{\mathrm{H}}\mathbf{P}_S \approx \mathbf{P}^{\mathrm{H}}\mathbf{P}, \tag{5.25}$$

as $\mathbf{P}_S^{\mathrm{H}}\mathbf{P}_S = (\mathbf{G}^{\mathrm{T}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^*)^{-1}$ and $\mathbf{P}^{\mathrm{H}}\mathbf{P} = (\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1}$. It is efficient to write $\mathbf{P}^{\mathrm{H}}$ as

$$\mathbf{P}^{\mathrm{H}} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \dots \mathbf{p}_M \end{bmatrix}$$

so that matrix $\mathbf{P}^{\mathrm{H}}\mathbf{P}$ can be understood as the summation of rank-one matrices,

$$\mathbf{P}^{\mathrm{H}}\mathbf{P} = \sum_{m=1}^{M} \mathbf{p}_m \mathbf{p}_m^{\mathrm{H}}, \tag{5.26}$$

and the approximation in equation (5.25) is rewritten as

$$\begin{aligned} \sum_{m=1}^{M} z_m \mathbf{p}_m \mathbf{p}_m^{\mathrm{H}} &\approx \sum_{m=1}^{M} \mathbf{p}_m \mathbf{p}_m^{\mathrm{H}} \\ \begin{bmatrix} \mathbf{p}_1\mathbf{p}_1^{\mathrm{H}} & \mathbf{p}_2\mathbf{p}_2^{\mathrm{H}} & \cdots \mathbf{p}_M\mathbf{p}_M^{\mathrm{H}} \end{bmatrix} \begin{bmatrix} z_1\mathbf{I}_K \\ z_2\mathbf{I}_K \\ \vdots \\ z_m\mathbf{I}_K \end{bmatrix} &\approx \sum_{m=1}^{M} \mathbf{p}_m \mathbf{p}_m^{\mathrm{H}} \end{aligned} \tag{5.27}$$

which is more convenient for MP purposes. Although the approximation in equation (5.27) is between two matrices and not two vectors as in the MP technique, we can get around this issue by following the same manipulations mentioned in the MR case. Hence, we are able to employ the MP method by setting $\mathbf{b} = \mathrm{vec}(\mathbf{P}^{\mathrm{H}}\mathbf{P})$ and

$$\mathbf{D}_{\mathrm{ZF}} = \begin{bmatrix} \mathbf{p}_1\mathbf{p}_1^{\mathrm{H}} & \mathbf{p}_2\mathbf{p}_2^{\mathrm{H}} & \cdots \mathbf{p}_M\mathbf{p}_M^{\mathrm{H}} \end{bmatrix} \tag{5.28}$$

as inputs in the MPAS algorithm in Algorithm 5. As a result we obtain the Algorithm 7.

---

**Algorithm 7** : Proposed ZF Matching Pursuit Antenna Selection (ZF-MPAS)

---

1) Input: $\mathbf{b} = \mathrm{vec}(\mathbf{P}^{\mathrm{H}}\mathbf{P})$, $\mathbf{D}_{\mathrm{ZF}} = [\mathbf{c}_1 \ \ldots \ \mathbf{c}_M]$, $\mathcal{I} = \{1, \ldots M\}$
2) Initialization: $i = 1$, $\mathbf{r}_1 = \mathbf{b}$, $\mathbf{z} = \mathbf{0}_M$
3) Repeat until $i = S - 1$:
a) Find the closest codeword, i.e., find $m_i \in \mathcal{I}$ such that

$$\langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle = \max_{i \in \mathcal{I}} \{ |\langle \mathbf{r}_i, \mathbf{c}_m \rangle| \}$$

b) Choose $z_{m_i} = 1$
c) Let $\mathbf{r}_{i+1} = \mathbf{r}_i - z_{m_i}\mathbf{c}_{m_i}$
d) $\mathcal{I} = \mathcal{I} - m_i$
e) Increment $i$.
4)Stop.
5) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

6) Output: $\mathbf{G}_S$

---

### 5.2.3 Symbol-level Matching Pursuit Antenna Selection

In the downlink of a massive MIMO system, the BS desires to transmit vectors of the form $\mathbf{q} \in \mathbb{C}^{(K\times 1)}$ containing the symbols intended to each terminal $k, k \in \{1, \cdots, K\}$. In fact, the BS generates a block of $L$ message vectors $[\mathbf{q}_1 \ \ldots \mathbf{q}_L]$. First, the BS performs precoding so that the transmitted signal is $\mathbf{x} = \mathbf{P}\mathbf{q}$, when all the transmit antennas are active. The precoding matrix $\mathbf{P}$ can be the same for all $\mathbf{q}_l$, in the transmit block. The preceding antenna selection methods based on MP compute the matrix $\mathbf{P}$ only once. Therefore, the associated precoders are called channel-level precoders, which means that the AS algorithms have as input only the channel matrix and the number of selected antennas. In contrast, matrix $\mathbf{P}$ can be computed for every $\mathbf{q}_l$ in the transmit block. In the present case, the AS algorithms have additionally the desired symbol vector $\mathbf{q}_l$ as input, and hence the associated precoder is known as symbol-level precoder. In this subsection, we propose an AS algorithm, also based on MP, in which the associated precoder is a symbol-level precoder.

We start by considering the message which is indeed transmitted by the $M$ BS antennas

$$\mathbf{P}\mathbf{q} = \mathbf{x} \tag{5.29}$$

and we left-multiply both sides of equation (5.29) by $\mathbf{P}^{\mathrm{H}}$ yielding

$$\mathbf{P}^{\mathrm{H}}\mathbf{P}\mathbf{q} = \mathbf{P}^{\mathrm{H}}\mathbf{x}. \tag{5.30}$$

Fortunately, we have just obtained a vector equal to a matrix multiplied by a vector, and thus we can bring it to the MP point of view. We only need to consider $\mathbf{x}$ as the selection vector $\mathbf{z}$, that is, we need to build $\mathbf{z}$ that reflects the most informative elements of $\mathbf{x}$. The resulting problem can be formulated as

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} \quad & \left\|\mathbf{P}^{\mathrm{H}}\mathbf{P}\mathbf{q} - \mathbf{P}^{\mathrm{H}}\mathbf{z}\right\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S, \quad \mathbf{z} \in \{0,1\}^M. \end{aligned} \tag{5.31}$$

If we employ a ZF precoder, we can replace matrix $\mathbf{P}$ by $\mathbf{P}_{\mathrm{ZF}} = \mathbf{G}^*(\mathbf{G}^{\mathrm{T}}\mathbf{G}^*)^{-1}$ in the problem represented by equation (5.31) and obtain

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} \quad & \left\|\mathbf{q} - \mathbf{G}^{\mathrm{T}}\mathbf{z}\right\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S, \quad \mathbf{z} \in \{0,1\}^M \end{aligned} \tag{5.32}$$

which is similar to the problem solved by the LASSO-AS algorithm [21] and can be solved by the Algorithm 8. On the other hand, if we opt for an MR precoding

scheme, problem (5.31) becomes

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} \quad & \left\|\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*}\mathbf{q}-\mathbf{G}^{\mathrm{T}}\mathbf{z}\right\|_2^2 \\ \text{subject to} \quad & \|\mathbf{z}\|_0 = S, \quad \mathbf{z}\in\{0,1\}^M \end{aligned} \tag{5.33}$$

and hence can be solved by the Algorithm 9. Interestingly, the simulation results showed that there is no loss in performance if we choose $\mathbf{b}=\mathbf{q}$ instead of $\mathbf{b}=\mathbf{P}^{\mathrm{H}}\mathbf{P}\mathbf{q}$. Hence the initial residue is set as $\mathbf{b}=\mathbf{q}$ in Algorithms 8 and 9.

**Algorithm 8** : Proposed Zero-Forcing Symbol-level Matching Pursuit Antenna Selection (ZFSL-MPAS)

1) Input: $\mathbf{b}=\mathbf{q}$, $\mathbf{D}=[\mathbf{c}_1 \ \dots \ \mathbf{c}_M]=\mathbf{P}_{\mathrm{ZF}}^{\mathrm{H}}$, $\mathcal{I}=\{1,\dots M\}$
2) Initialization: $i=1$, $\mathbf{r}_1=\mathbf{b}$, $\mathbf{z}=\mathbf{0}_M$
3) Repeat until $i=S-1$:
a) Find the closest codeword, i.e., find $m_i\in\mathcal{I}$ such that

$$\langle \mathbf{r}_i,\mathbf{c}_{m_i}\rangle = \max_{i\in\mathcal{I}} \{|\langle \mathbf{r}_i,\mathbf{c}_m\rangle|\}$$

b) Choose $z_{m_i}=1$
c) Let $\mathbf{r}_{i+1}=\mathbf{r}_i-z_{m_i}\mathbf{c}_{m_i}$
d) $\mathcal{I}=\mathcal{I}-m_i$
e) Increment $i$.
4)Stop.
5) Compute the $S$-selected channel matrix

$$\mathbf{G}_S=\mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G}=\mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

6) Output: $\mathbf{G}_S$

## 5.3 Concluding Remarks

The main contribution of this chapter was two antenna selection algorithms based on the MP technique. In fact, since we are considering two precoding schemes, each algorithm has two versions depending on the chosen precoding approach. The MPAS algorithm is associated with a channel-level precoder, whereas the SL-MPAS should work along with a symbol-level precoder. Symbol-level precoders are computationally more costly than channel-level precoders. Nonetheless, the symbol-level precoders achieve the lowest bit error rate, as we will see in the simulation results in chapter 6. In the next chapter, the performance of each AS algorithm is evaluated in terms of bit error rate. Moreover, the number of operations of our proposed AS methods are compared with the corresponding AS algorithm counterparts.

**Algorithm 9** : Proposed Symbol-level Matching Pursuit Antenna Selection (SL-MPAS)

1) Input: $\mathbf{b} = \mathbf{q}$, $\mathbf{D} = [\mathbf{c}_1 \ \ldots \ \mathbf{c}_M] = \mathbf{P}_{\mathrm{MR}}^{\mathrm{H}}$, $\mathcal{I} = \{1, \ldots M\}$
2) Initialization: $i = 1$, $\mathbf{r}_1 = \mathbf{b}$, $\mathbf{z} = \mathbf{0}_M$
3) Repeat until $i = S - 1$:
a) Find the closest codeword, i.e., find $m_i \in \mathcal{I}$ such that

$$\langle \mathbf{r}_i, \mathbf{c}_{m_i} \rangle = \max_{i \in \mathcal{I}} \{ |\langle \mathbf{r}_i, \mathbf{c}_m \rangle| \}$$

b) Choose $z_{m_i} = 1$
c) Let $\mathbf{r}_{i+1} = \mathbf{r}_i - z_{m_i} \mathbf{c}_{m_i}$
d) $\mathcal{I} = \mathcal{I} - m_i$
e) Increment $i$.
4)Stop.
5) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

6) Output: $\mathbf{G}_S$

# Chapter 6

# Simulation Results

In this chapter, the proposed antenna selection algorithms, introduced in chapter 5, are evaluated via simulation. We also provide the comparison among the proposed algorithms and the existent AS algorithms presented in chapter 4 for massive MIMO system. The comparison is performed via bit error rate (BER). We assume perfect and imperfect channel estimation at BS. All the simulation results are obtained by averaging 300 Monte Carlo runs. The simulations are performed using MATLAB software and the convex optimization problems are solved using CVX [102].

Since the primary motivation for the proposed AS algorithms is their reduced computational complexity, in this chapter we also quantify their complexity by counting the required number of flops to compute the selection vector $\mathbf{z}$. A flop is defined to be a real floating point operation [100]. A real addition or multiplication is counted as one flop, whereas a complex addition and multiplication have two flops and six flops, respectively [101]. We also provide the number of flops required to compute the transmit vector $\mathbf{x}$ in a massive MIMO system.

To compare our proposed AS algorithms with their counterparts, we provide the time spent by each AS algorithm to compute a block of 50 transmitted vectors $\mathbf{x}$ in a massive MIMO system. The simulations are conducted in a computer with Intel Core i7-7500U CPU 2.70GHz x4 processor and 7.7 GB of memory.

## 6.1 BER simulations

In this section, we investigate how selecting a subset of $S$ active BS antennas out of $M$ can impact the average BER per user performance, as the SNR varies in the downlink of a massive MIMO system. The system is composed of $K = 12$ single-antenna terminals that are served by a BS equipped with $M = 100$ antennas in a single-cell.

The simulation steps are outlined in Table 6.1. First of all, the BS uses the pilot

symbols, transmitted by the $K$ terminals in the uplink, to obtain an estimate of the channel matrix. Secondly, the BS generates a block of $L = 50$ message vectors $[\mathbf{q}_1 \ \ldots \mathbf{q}_{50}]$, where each vector $\mathbf{q}$ contains a stream of BPSK or 4-QAM samples. Thirdly, the BS applies the AS algorithms which are listed in Table 6.2, where the proposed AS algorithms are highlighted in boldface. In the symbol-level precoder case, steps 4-5 in Table 6.1 are repeated whenever a vector $\mathbf{q}$ is generated. In contrast, steps 4-5 are performed only once if the associated precoder is channel-level. All the AS algorithms generate the selected channel matrix $\mathbf{G}_S$, except the LASSO-AS algorithm which generates the transmitted signal $\mathbf{x}$ directly. Therefore, in the LASSO-AS case, steps 5 and 6 are not performed. We consider both uncorrelated and correlated Rayleigh fading channel models which are detailed in equations (3.4) and (3.8), respectively. We also consider perfect and imperfect CSI knowledge at the BS.

The massive MIMO simulation results comprise curves of the Average BER per user for $S = \{34, 91\}$ when the SNR is varied between -12 and 12 dB. In these results, two curves are always present in order to provide upper and lower bounds. The first one is composed of the ZF-full and MR-full methods, in which all the BS antennas are active and hence we expect to be the best BER performance. The second one is composed of the ZF-RAS and MR-RAS methods, in which the AS algorithm is random and then we expect to be the worst BER performance. It is worth mentioning that we set the sparsity factor $\alpha_1 = \{0.5, 0.9\}$ in the LASSO algorithm to obtain $S = \{34, 91\}$, as recommended in [21].

Steps 1, 3 and 5 contain two possible choices related to channel model, constellation-symbol types and precoding scheme, respectively. These options are used to organize the massive MIMO simulation results by scenarios.

In scenario 1, the channel is modeled assuming a rich scattering environment, which is equivalent to using an uncorrelated Rayleigh fading model. The transmitted message $\mathbf{x}$ is composed of BPSK symbols, and both ZF and MR precoding schemes are employed. Scenario 2 assumes the same channel model considered in scenario 1. These scenarios differ with respect to the constellation symbol, which is 4-QAM for scenario 2 and BPSK for scenario 1.

In scenario 3, the channel is modeled assuming a poor scattering environment [61] in which only 2 paths reach the receivers, that is equivalent to use a correlated Rayleigh fading model with $N = 2$ in equation (3.8). As in scenario 1, the transmitted signal is composed by BPSK symbols.

Scenario 4 is a version of scenario 3 with 4-QAM transmitted symbols.

It is worth to mentioning that the BS has perfect CSI knowledge in the four previous scenarios. On the other hand, scenarios 5-8 are equivalent to scenarios 1-4 with imperfect CSI knowledge at the BS. Furthermore, each scenario has two

Table 6.1: Simulation steps (massive MIMO)

1) Generate the colums of the channel matrix $\mathbf{G}$

   a) $\mathbf{g}_k \leftarrow \begin{cases} \text{realization of } \boldsymbol{g} \sim \mathcal{CN}(\mathbf{0}_{(M)},\, \beta_k \mathbf{I}_{(M)}) \\ \sum_{n=1}^{N} h_n \left[1 \;\; e^{2\pi j d_H \sin(\theta_n)} \ldots e^{2\pi j d_H (M-1)\sin(\theta_n)}\right]^{\mathrm{T}} \end{cases}$

   b) Normalize the columns of $\mathbf{G}$

2) Perform channel estimation at BS (with Uplink SNR $\rho_{\mathrm{ul}}$)

   a) The BS receives the pilot signals transmitted by the $K$ terminals

   $\mathbf{Y}_{\mathrm{p}} = \mathbf{G}\mathbf{X}_{\mathrm{p}} + \mathbf{W}_{\mathrm{p}} = \sqrt{\tau_{\mathrm{p}}}\mathbf{G}\mathbf{\Phi}^{\mathrm{H}} + \mathbf{W}_{\mathrm{p}}$

   b) Perform De-spreading by right-multiplying by the pilot matrix $\mathbf{\Phi}$

   $\mathbf{Y}_{\mathrm{p}} = \sqrt{\tau_{\mathrm{p}}}\mathbf{G}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi} + \mathbf{W}_{\mathrm{p}}\mathbf{\Phi} = \sqrt{\tau_{\mathrm{p}}}\mathbf{G} + \mathbf{W}_{\mathrm{p}}\mathbf{\Phi}$

   c) Obtain a channel estimate by using an MMSE estimator

   $\widehat{\mathbf{G}} = \mathrm{MMSE}(\mathbf{Y}_{\mathrm{p}})$

3) Generate the intended message for each terminal

   a) $\mathbf{q} \leftarrow \begin{cases} \text{BPSK symbols} \\ \text{4-QAM symbols} \end{cases}$

4) BS decides which antennas to transmit

   a) Apply the AS algorithms (summarized in Table 6.2) to obtain $\mathbf{G}_S$

5) Generate normalized precoding matrix

   a) $\mathbf{P}_S \leftarrow \begin{cases} \mathbf{G}_S^*(\mathbf{G}_S^{\mathrm{T}}\mathbf{G}_S^*)^{-1}, & \text{for ZF precoding} \\ \mathbf{G}_S^*, & \text{for MR precoding} \end{cases}$

6) Form the signal to transmit

   $\mathbf{x} = \mathbf{P}_S \mathrm{diag}\,(\eta)^{1/2}\,\mathbf{q}$

7) Perform signal reception at each terminal

   a) Each terminal $k$ receives

   $y_k = \mathbf{g}_k^T\mathbf{x} + w_k$

8) Calculate average BER per user,

   $\overline{\mathrm{BER}} = \frac{1}{K}\sum_{k=1}^{K} \mathrm{BER}_k$,

   where $\mathrm{BER}_k$ for $k \in \{1, \cdots, K\}$ is the BER of user $k$

9) Compute the average SNR per user

   $\rho_{\mathrm{dl}} = 10\ \log_{10}\left(\frac{\mathrm{Pow}_T}{\mathrm{Pow}_R}\right)$

   where $\mathrm{Pow}_T$ and $\mathrm{Pow}_R$ are the average transmitting and receiving power, respectively.

Table 6.2: Antenna selection algorithms evaluated in the BER simulations

| **Algorithm** | **Description** | **Precoder** | **Location** |
|---|---|---|---|
| **ZF-MPAS** | Zero forcing - Matching Pursuits Antenna Selection | channel-level | Algorithm 7 |
| RAS | Random Antenna Selection | channel-level | Algorithm 3 |
| MCAS [20] | Maximum Capacity Antenna Selection | channel-level | Algorithm 1 |
| **ZFSL-MPAS** | Zero forcing Symbol Level - Matching Pursuits Antenna Selection | symbol-level | Algorithm 8 |
| LASSO [21] | Least Absolute Shrinkage and Selection Operator | symbol-level | Algorithm 2 |
| **MR-MPAS** | Maximum Ratio - Matching Pursuits Antenna Selection | channel-level | Algorithm 6 |
| **MRSL-MPAS** | Maximum Ratio Symbol Level - Matching Pursuits Antenna Selection | symbol-level | Algorithm 9 |

groups of results: the ones achieved with ZF precoding and the ones achieved with MR precoding.

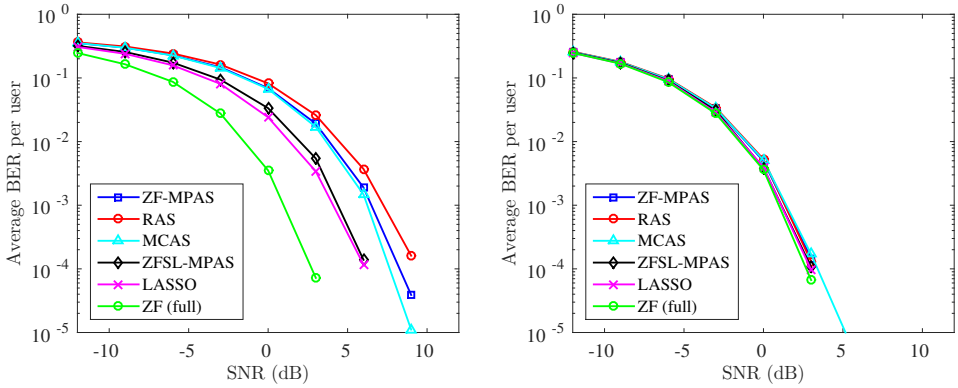
### 6.1.1 Perfect CSI knowledge at BS

In this subsection, the channel estimation performed by the BS in the uplink is considered perfect, that is, the uplink SNR is assumed high, for example $\rho_{\mathrm{dl}} = 12\,dB$. Hence, the channel estimate $\widehat{\mathbf{G}}$ is equal to the actual channel matrix $\mathbf{G}$. Figure 6.2 depicts the Average BER per user for scenario 1. For the first four scenarios with ZF precoding scheme, illustrated in Figures 6.2(a,b), 6.4(a,b), 6.6(a,b) and 6.8(a,b), the BER curves are slightly different. First due to the constellation chosen, i.e., BPSK (in Figure 6.2(a,b)) achieves better performance than 4-QAM (in Figure 6.4(a,b)). Moreover a poor scattering environment induces a worse performance (in Figure 6.6(a,b)), when compared with a rich scattering one (in Figure 6.2(a,b)). Nevertheless, the AS algorithms tested in these scenarios (with ZF precoding) perform quite the same. When only 34 antennas are active, the proposed ZFSL-MPAS algorithm is always very close in performance to the LASSO algorithm, with the benefit of being less computationally expensive. The same holds for the proposed ZF-MPAS algorithm which is always quite close in performance to the MCAS algorithm, but also has the advantage of having an associated channel-level precoder. As expected, when 91% of the BS antennas are active, the AS methods behavior

is very similar to the one when all the antennas are active. This also uncovers the massive MIMO ability to operate well even when 9% of the antennas are damaged [21].

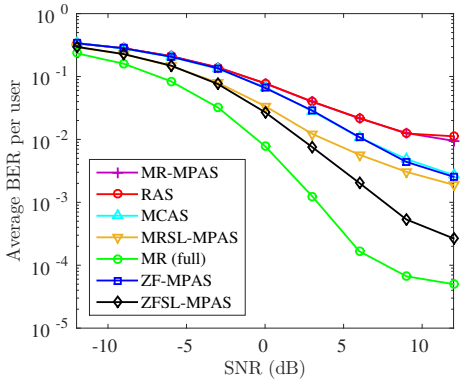
On the other hand, we can note that the overall performance is harmed when the MR precoding is chosen. In fact, MR precoding relies on favorable propagation, i.e., $\mathbf{G}^{\mathrm{T}}\mathbf{G}^{*} = \mathbf{I}_M$, which might not be true in the simulations, specially in the scenario with too few paths reaching the terminals.

Since the MR-MPAS and MRSL-MPAS algorithms performance was not good, we also employed the ZF-MPAS and ZFSL-MPAS algorithms, that were originally conceived for ZF precoding, with MR precoding. In this way, ZF-MPAS and ZFSL-MPAS algorithms select the subset of antennas pretending that a ZF precoding is going to be applied after, but in fact an MR precoding is used in the end. Interestingly, the results are quite promising, as one can see in Figures 6.2(c,d), 6.4(c,d), 6.6(c,d) and 6.8(c,d) but specially in Figure 6.6(c,d), where ZFSL-MPAS with 34 active antennas outperforms MR-full with all active antennas. One possible reason is that by activating all the antennas the propagation becomes less favorable when there is poor scattering in the environment.

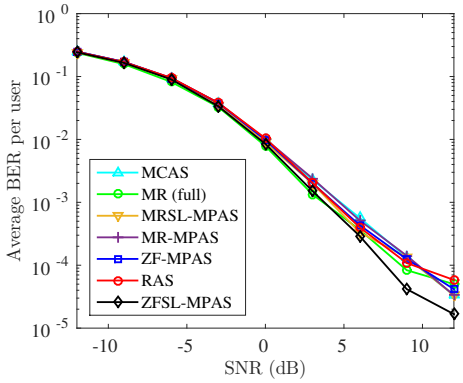



(a) ZF precoding scheme, $S = 34$. (b) ZF precoding scheme, $S = 91$.

Figure 6.1: ZF precoding and Scenario 1: Average BER per user for a massive MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols.

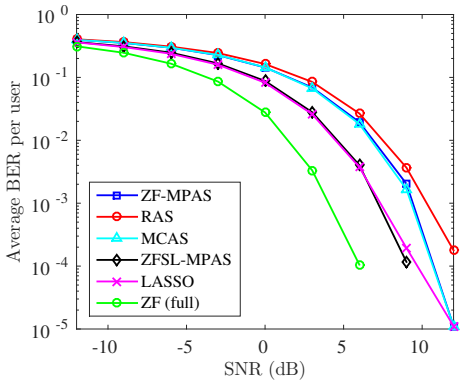



(a) MR precoding scheme, $S = 34$.

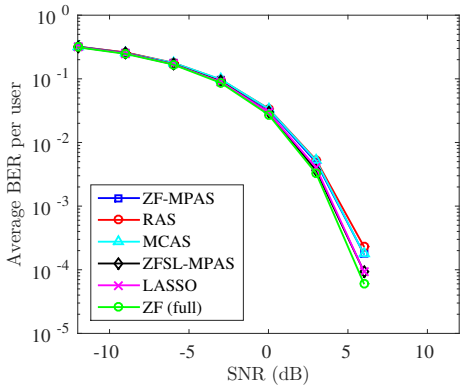



(b) MR precoding scheme, $S = 91$.

Figure 6.2: MR precoding Scenario 1: Average BER per user for a massive MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols.

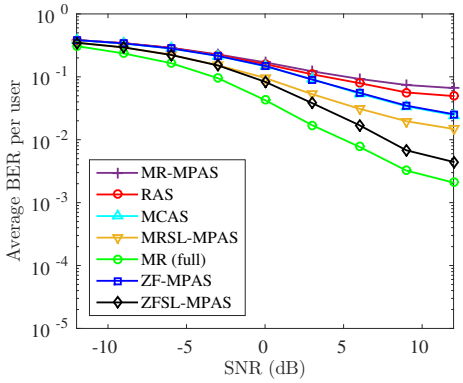



(a) ZF precoding scheme, $S = 34$.

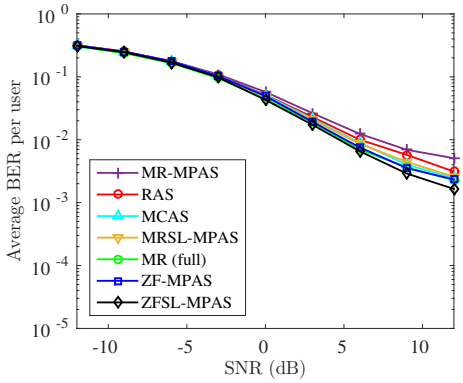



(b) ZF precoding scheme, $S = 91$.

Figure 6.3: ZF precoding and Scenario 2: Average BER per user for a massive MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and 4-QAM transmitted symbols.

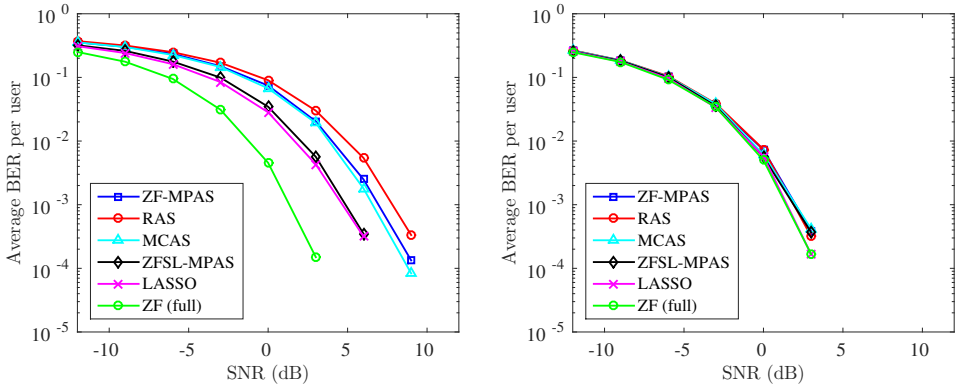



(a) MR precoding scheme, $S = 34$.



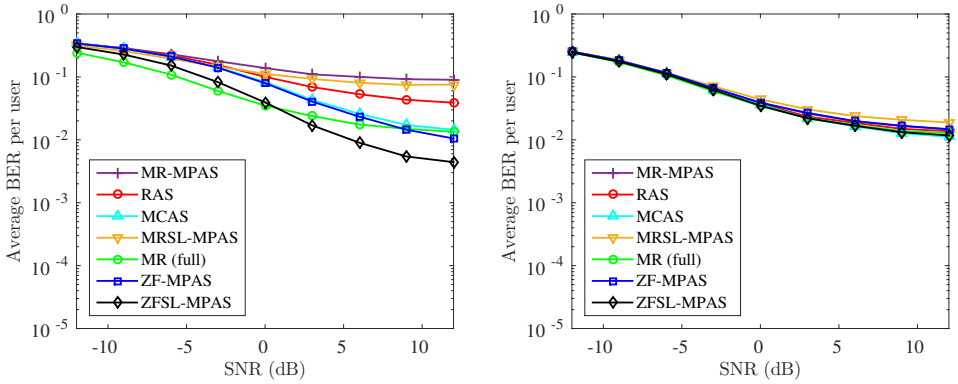

(b) MR precoding scheme, $S = 91$.

Figure 6.4: MR precoding and Scenario 2: Average BER per user for a massive MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and 4-QAM transmitted symbols.



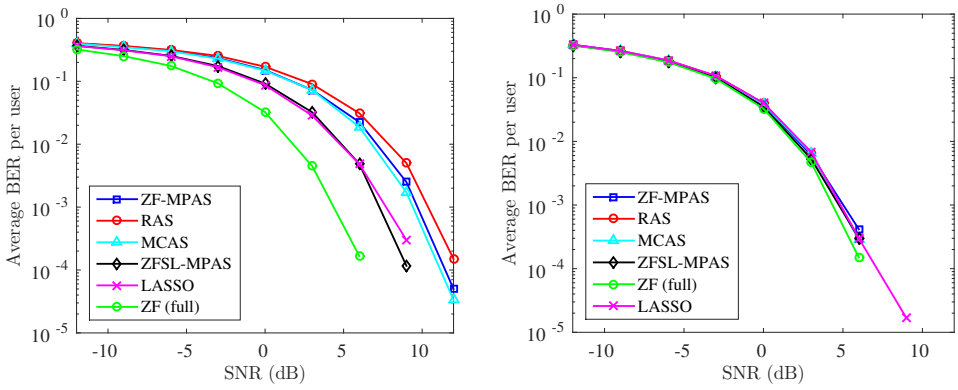

(a) ZF precoding scheme, $S = 34$. (b) ZF precoding scheme, $S = 91$.

Figure 6.5: ZF precoding and Scenario 3: Average BER per user for a massive MIMO system with perfect CSI knowledge, correlated Rayleigh channel, and BPSK transmitted symbols.



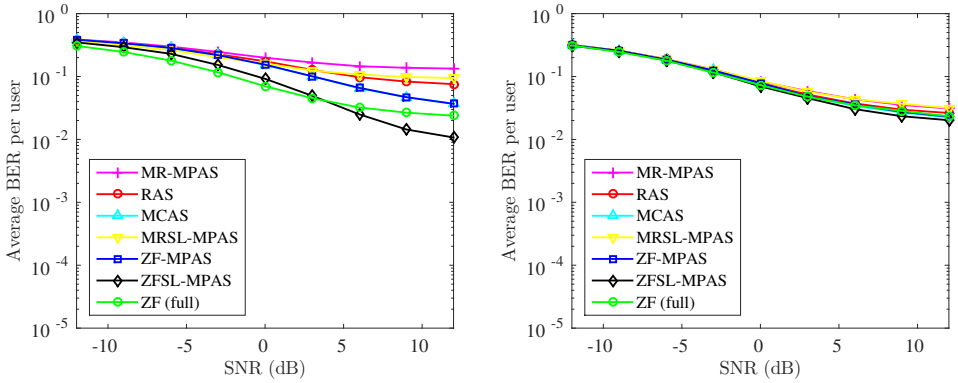

(a) MR precoding scheme, $S = 34$. (b) MR precoding scheme, $S = 91$.

Figure 6.6: MR precoding and Scenario 3: Average BER per user for a massive MIMO system with perfect CSI knowledge, correlated Rayleigh channel, and BPSK transmitted symbols.




(a) ZF precoding scheme, $S = 34$. (b) ZF precoding scheme, $S = 91$.

Figure 6.7: ZF precoding and Scenario 4: Average BER per user for a massive MIMO system with perfect CSI knowledge, correlated Rayleigh channel, and 4-QAM transmitted symbols.



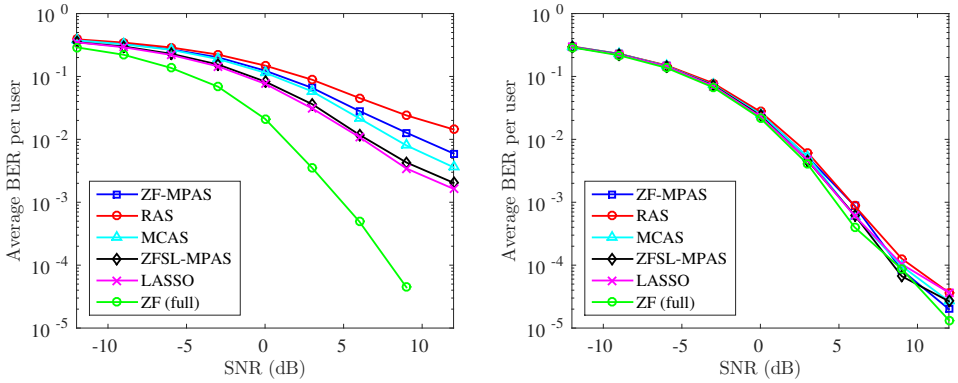

(a) MR precoding scheme, $S = 34$. (b) MR precoding scheme, $S = 91$.

Figure 6.8: MR precoding and Scenario 4: Average BER per user for a massive MIMO system with perfect CSI knowledge, correlated Rayleigh channel, and 4-QAM transmitted symbols.

### 6.1.2 Imperfect CSI knowledge at BS

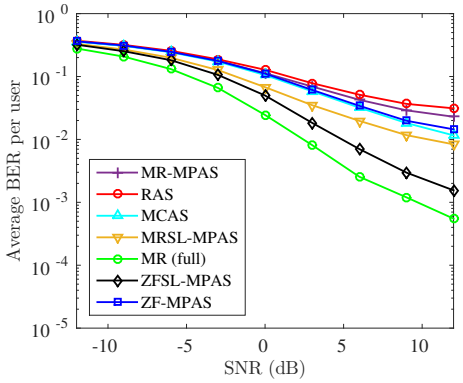
In this scenario, we investigate the impact on the results of not having a very accurate channel estimate. Then, we decrease the Uplink SNR $\rho_{\text{dl}}$ to 3 dB. The channel estimate is the output of the MMSE estimator.

In the last four scenarios with ZF precoding scheme, depicted in Figures 6.10(a,b), 6.12(a,b), 6.14(a,b) and 6.16(a,b), the patterns formed by the BER curves are very similar to the ones achieved by the MR precoding in the previous simulations. Thus, when the BS has poor CSI knowledge we should not expect measured BER equal to zero at high SNR. For the cases where an MR precoder is employed, the results are slightly worse than the ones yielded when the BS knows the CSI perfectly.

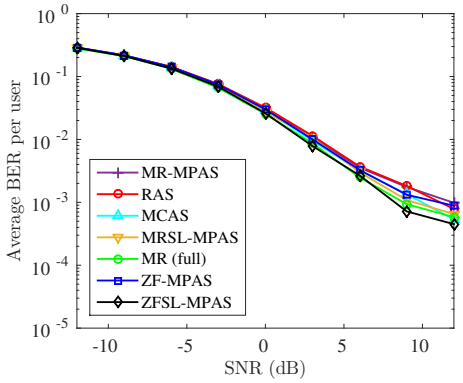



(a) ZF precoding scheme, $S = 34$. (b) ZF precoding scheme, $S = 91$.

Figure 6.9: ZF precoding and Scenario 5: Average BER per user for a massive MIMO system with imperfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols.

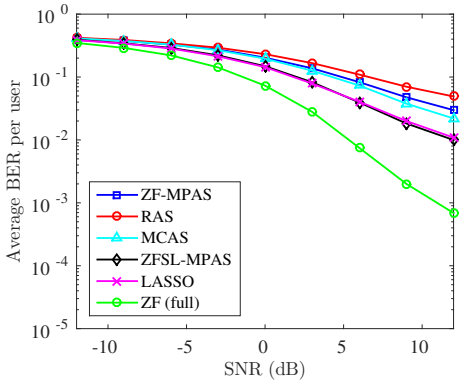



(a) MR precoding scheme, $S = 34$.

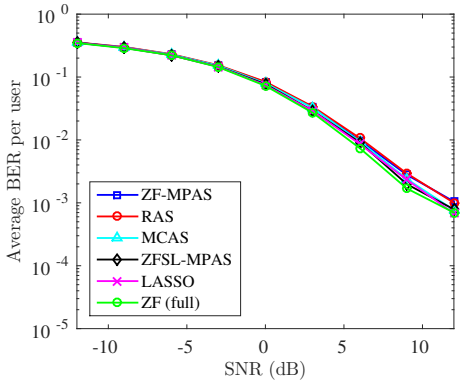



(b) MR precoding scheme, $S = 91$.

Figure 6.10: MR precoding and Scenario 5: Average BER per user for a massive MIMO system with imperfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols.

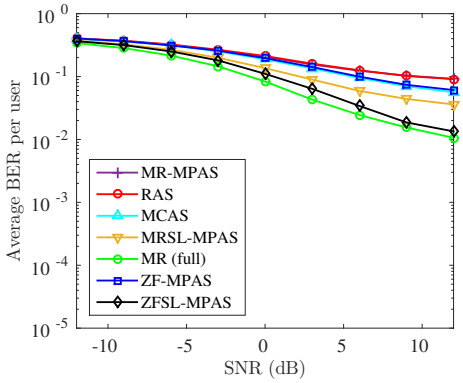



(a) ZF precoding scheme, $S = 34$.

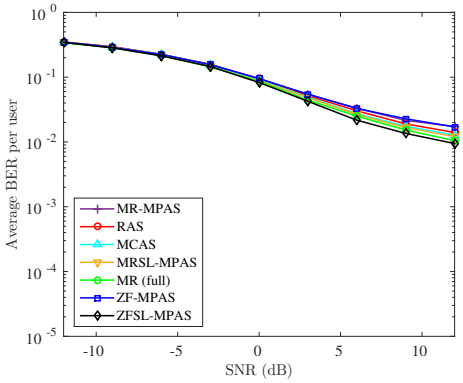



(b) ZF precoding scheme, $S = 91$.

Figure 6.11: ZF precoding and Scenario 6: Average BER per user for a massive MIMO system with imperfect CSI knowledge, uncorrelated Rayleigh channel, and 4-QAM transmitted symbols.

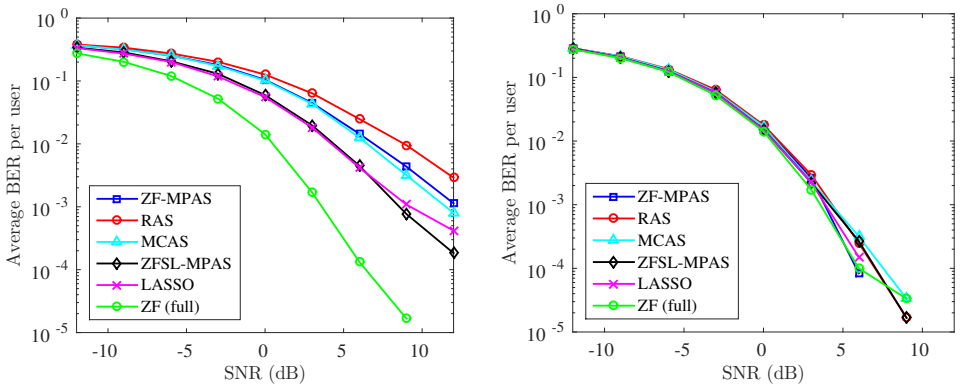



(a) MR precoding scheme, $S = 34$.



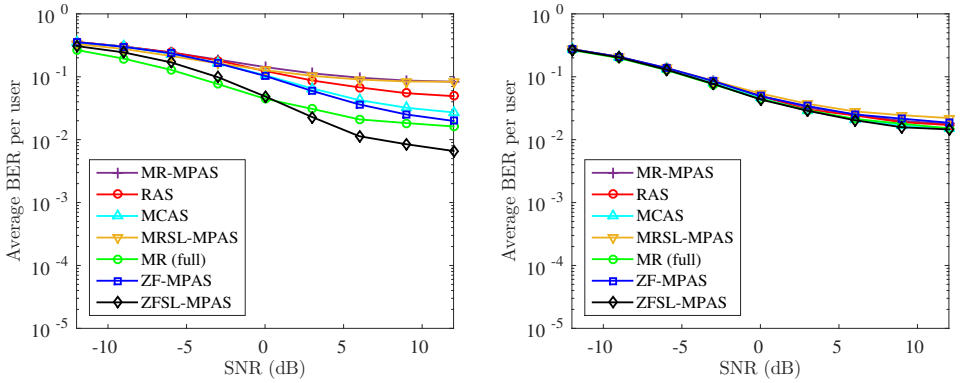

(b) MR precoding scheme, $S = 91$.

Figure 6.12: MR precoding and Scenario 6: Average BER per user for a massive MIMO system with imperfect CSI knowledge, uncorrelated Rayleigh channel, and 4-QAM transmitted symbols.



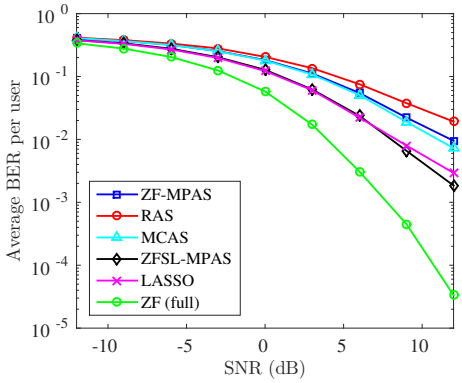

(a) ZF precoding scheme, $S = 34$. (b) ZF precoding scheme, $S = 91$.

Figure 6.13: ZF precoding and Scenario 7: Average BER per user for a massive MIMO system with imperfect CSI knowledge, correlated Rayleigh channel, and BPSK transmitted symbols.

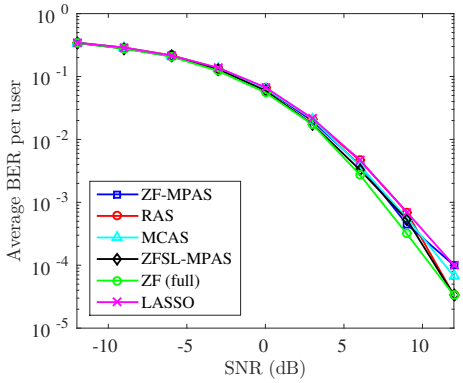



(a) MR precoding scheme, $S = 34$. (b) MR precoding scheme, $S = 91$.

Figure 6.14: MR precoding and Scenario 7: Average BER per user for a massive MIMO system with imperfect CSI knowledge, correlated Rayleigh channel, and BPSK transmitted symbols.

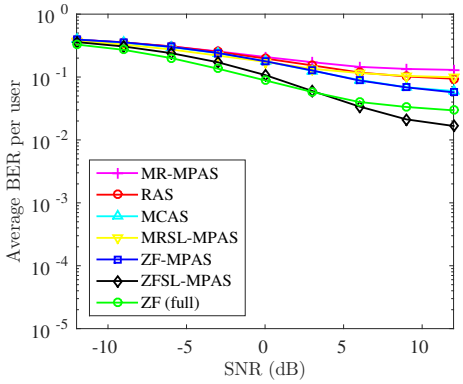



(a) ZF precoding scheme, $S = 34$.

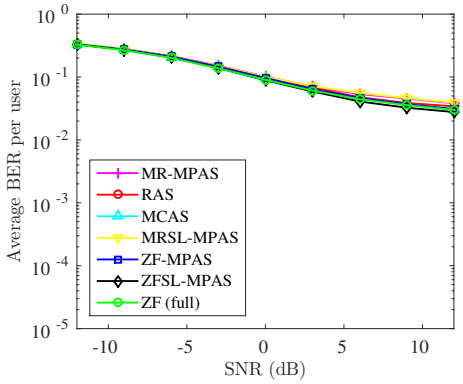



(b) ZF precoding scheme, $S = 91$.

Figure 6.15: ZF precoding and Scenario 8: Average BER per user for a massive MIMO system with imperfect CSI knowledge, correlated Rayleigh channel, and 4-QAM transmitted symbols.




(a) MR precoding scheme, $S = 34$.




(b) MR precoding scheme, $S = 91$.

Figure 6.16: MR precoding and Scenario 8: Average BER per user for a massive MIMO system with imperfect CSI knowledge, correlated Rayleigh channel, and 4-QAM transmitted symbols.

## 6.2 Computational Complexity

### 6.2.1 Common Matrix Operations

In this subsection, we summarize the most common matrix operations used to compute the computational cost of each AS algorithm.

#### 6.2.1.1 Inner product

The inner product between two vectors of size $n$ involves $n-1$ additions and $n$ multiplications. To simplify we consider that the inner product involves $n$ additions as considered in [100]. If the vectors are complex, then the dot product requires $4n$ additions and $4n$ multiplications and hence $8n$ flops.

#### 6.2.1.2 Outer product

The outer product between two vectors of size $n$ involves $n^2$ multiplications [100]. If the vectors are complex, then it requires $2n^2$ additions and $4n^2$ multiplications and hence $6n^2$ flops.

#### 6.2.1.3 Matrix-vector product

The product between a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ and a vector $\mathbf{b} \in \mathbb{C}^{n\times 1}$ is composed by $n$ inner products between vectors of size $n$. Therefore, it requires $8n^2$ flops.

#### 6.2.1.4 Matrix-matrix product

The product between two matrices $\mathbf{A} \in \mathbb{C}^{m\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times k}$ consist of $mk$ inner products between vectors of size $n$. Then, the resulting complexity is $mk(8n)$ flops. If we are interested in square matrices, that is $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$, then the matrix product involves $8n^3$ flops.

#### 6.2.1.5 Inverse of a square matrix

The matrix inversions are computed using QR decomposition, i.e. by computing $\mathbf{A} = \mathbf{QR}$, $\mathbf{A} \in \mathbb{C}^{n\times n}$. Then

$$\mathbf{A}^{-1} = \mathbf{R}^{-1}\mathbf{Q}^{\mathrm{H}} \tag{6.1}$$

where $\mathbf{R} \in \mathbb{C}^{n\times n}$ is upper triangular with real nonzero diagonal elements and $\mathbf{Q} \in \mathbb{C}^{n\times n}$ has orthonormal columns ($\mathbf{Q}^{\mathrm{H}}\mathbf{Q} = \mathbf{I}_n$). The complex QR decomposition requires $16n^3/3$ flops [100]. If we left-multiply equation (6.1) by $\mathbf{R}$ it yields

$$\mathbf{R}\mathbf{A}^{-1} = \mathbf{Q}^{\mathrm{H}}, \tag{6.2}$$

which can be seen as $n$ linear systems if we rearrange equation (6.2) as

$$\mathbf{R}\begin{bmatrix}\widetilde{\mathbf{a}}_1 & \widetilde{\mathbf{a}}_2 & \dots & \widetilde{\mathbf{a}}_n\end{bmatrix} = \begin{bmatrix}\widetilde{\mathbf{q}}_1 & \widetilde{\mathbf{q}}_2 & \dots & \widetilde{\mathbf{q}}_n\end{bmatrix}, \tag{6.3}$$

in which $\widetilde{\mathbf{a}}_i$ and $\widetilde{\mathbf{q}}_i$ are the $i$th column of $\mathbf{A}^{-1}$ and $\mathbf{Q}^{\mathrm{H}}$, respectively. Then we can use Backward Substitution to solve each linear system in equation (6.3), which requires $4n^3 - 3n^2$ flops [100] in total. Therefore, the resulting complexity for inverting a matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ is $28n^3/3 - 3n^2$ flops.

### 6.2.2 Computing selection vector

In this subsection we count the total number of flops to produce the selection vector $\mathbf{z}$ in each proposed antenna selection algorithm. The proposed antenna selec-

tion algorithms are highlighted in boldface in Table 6.2. The complexity (in flops) of each proposed AS algorithm is summarized in Table 6.3.

#### 6.2.2.1 ZF-MPAS

First we need to compute the initial residue

$$\mathbf{b} = \text{vec}(\mathbf{P}^{\text{H}}\mathbf{P}) = \text{vec}((\mathbf{G}^{\text{T}}\mathbf{G}^{*})^{-1}), \tag{6.4}$$

where $\mathbf{P} = \mathbf{G}^{*}(\mathbf{G}^{\text{T}}\mathbf{G}^{*})^{-1}$ is the ZF precoding matrix. It requires $28K^3/3 - 3K^2 + 8M^2K$ flops, according to subsection 6.2.1.5. Secondly, we need to compute the dictionary matrix $\mathbf{D}$, which involves $M$ outer-products that require $6MK^2$ flops in total.

For each iteration we need to compute one addition, which requires $2K^2$ flops. Since $S$ iterations are performed in the ZF-MPAS algorithm, it results in $2SK^2$ flops. As the number of inner products that need to be computed in the search phase is decreased by one whenever an iteration is concluded, the number of inner products in $S$ iterations is $M + M - 1 + M - 2 + \ldots + M - (S - 1) = (MS - (S^2 - S)/2)$. To compute $(MS - (S^2 - S)/2)$ inner products of vectors of size $K^2$ it requires $(MS - (S^2 - S)/2)8K^2$ flops.

In total, the ZF-MPAS algorithm requires $28K^3/3 + 8M^2K + 2SK^2 + 8MSK^2 + 4SK^2 - 4S^2K^2 - 3K^2$ flops to produce the selection vector $\mathbf{z}$.

#### 6.2.2.2 MR-MPAS

First we need to compute the initial residue

$$\mathbf{b} = \text{vec}(\mathbf{P}^{\text{H}}\mathbf{P}) = \text{vec}(\mathbf{G}^{\text{T}}\mathbf{G}^{*}), \tag{6.5}$$

where $\mathbf{P} = \mathbf{G}^{*}$ is the MR precoding matrix. It involves $M$ additions and $M$ outer products, which requires $8MK^2$. Secondly, we need to compute the dictionary matrix $\mathbf{D}$, which involves $M$ outer-products but they were already counted in the previous step.

Similarly to ZF-MPAS, the $S$ iterations require $(MS - (S^2 - S)/2)8K^2 + 2SK^2$ flops. In total, the MR-MPAS algorithm requires $8MK^2 + (MS - (S^2 - S)/2)8K^2 + 2SK^2$ flops to obtain the selection vector $\mathbf{z}$.

#### 6.2.2.3 ZFSL-MPAS

First we need to compute the ZF precoding matrix $\mathbf{P} = \mathbf{G}^{*}(\mathbf{G}^{\text{T}}\mathbf{G}^{*})^{-1}$ which requires $28K^3/3 - 3K^2 + 8M^2K + 8MK^2$ flops. For each iteration we need to compute one addition, which requires $2K$ flops. Since $S$ iterations are performed in

the ZFSL-MPAS algorithm, it results in $2SK$ flops. To compute $(MS-(S^2-S)/2)$ inner products of vectors of size $K$ it requires $(MS-(S^2-S)/2)8K$ flops. Therefore, the ZFSL-MPAS algorithm requires $28K^3/3 - 3K^2 + 8M^2K + 8MK^2 + 6SK + 8MSK - 4S^2K$ flops to build the selection vector $\mathbf{z}$.

#### 6.2.2.4 MRSL-MPAS

The MR precoding matrix is $\mathbf{P} = \mathbf{G}^*$. In total $S$ iterations are required in the MRSL-MPAS algorithm, in which $(MS-(S^2-S)/2)8K+2SK$ flops are necessary. Therefore, the MRSL-MPAS algorithm requires $(MS-(S^2-S)/2)8K+2SK$ flops to produce the selection vector $\mathbf{z}$.

Table 6.3: Number of flops required to compute the selection vector by the Antenna selection algorithms highlighted in boldface in Table 6.2; $M$ is the number of BS antennas, $S$ is the number of selected antennas and $K$ is the number of terminals

| **AS Algorithm** | **Complexity (in flops)** |
|---|---|
| ZF-MPAS | $28K^3/3 + 8M^2K + 2SK^2 + 8MSK^2 + 4SK^2 - 4S^2K^2 - 3K^2$ |
| MR-MPAS | $8MK^2 + (MS-(S^2-S)/2)8K^2 + 2SK^2$ |
| ZFSL-MPAS | $28K^3/3 - 3K^2 + 8M^2K + 8MK^2 + 6SK + 8MSK - 4S^2K$ |
| MRSL-MPAS | $(MS-(S^2-S)/2)8K + 2SK$ |

### 6.2.3 Computing the vector to be transmitted

In this subsection we provide the number of flops required to compute the vector to be transmitted $\mathbf{x}$ for $K$ terminals in a massive MIMO system. The number of flops to compute the selection vector $\mathbf{z}$ was already computed in the previous section. After acquiring the selection vector $\mathbf{z}$, we can compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G} \tag{6.6}$$

which requires $8MSK$ flops (ignoring the flops needed to compute $\mathbf{z}$). To form the transmit message $\mathbf{x} = \mathbf{P}_S\mathbf{q}$ we first need to choose the precoding scheme.

- **ZF precoding scheme**
  First we compute the precoding matrix

$$\mathbf{P}_S = \mathbf{G}_S^*(\mathbf{G}_S^{\mathrm{T}}\mathbf{G}_S^*)^{-1} \tag{6.7}$$

  which requires $28K^3/3 - 3K^2 + 8SK^2$ flops. Then, to compute $\mathbf{x} = \mathbf{P}_S\mathbf{q}$ it is required $28K^3/3 - 3K^2 + 8SK^2 + 8MSK + 8SK$ flops. Therefore, each AS

algorithm in section 6.2.2 requires $28K^3/3 - 3K^2 + 8SK^2 + 8MSK + 8SK$ additional flops to build the transmit vector. The total number of flops is shown in Table 6.4.

- **MR precoding scheme**
  First we compute the precoding matrix

$$\mathbf{P}_S = \mathbf{G}_S^* \tag{6.8}$$

  which requires no flops. Then, to compute $\mathbf{x} = \mathbf{P}_S\mathbf{q}$ it is required $8MSK + 8SK$ flops. Therefore, each AS algorithm in section 6.2.2 requires $8MSK + 8SK$ additional flops to build the transmit vector. The total number of flops is shown in Table 6.5.

Figure 6.17 depicts the time spent by each AS algorithm to form a block of 50 transmit messages $\mathbf{x}$ in the downlink of a massive MIMO system. By comparing the two AS algorithms whose precoders are symbol-level in Figure 6.17, we can easily note that the proposed ZFSL-MPAS is less computationally intensive than LASSO. Moreover, we can observe that a symbol-level precoder requires less computation than a channel-level one, such as MCAS.

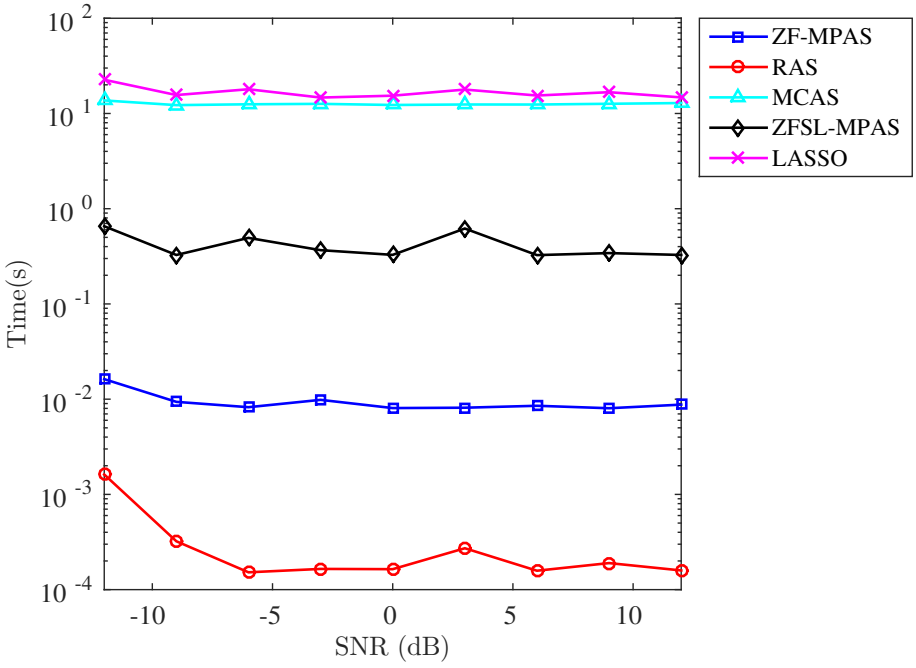



Figure 6.17: Time spent to compute 50 transmit messages by each AS algorithm in an massive MIMO system.

Table 6.4: Number of flops required to compute the message vector in massive MIMO (ZF precoding) by the proposed antenna selection algorithms highlighted in boldface in Table 6.2; $M$ is the number of BS antennas, $S$ is the number of selected antennas and $K$ is the number of terminals

| **AS Algorithm** | **Complexity (in flops)** |
|---|---|
| ZF-MPAS | $56K^3/3 + 8M^2K + 2SK^2 + 8MSK^2 +$ $12SK^2 - 4S^2K^2 - 6K^2 + 8MSK + 8SK$ |
| MR-MPAS | $8MK^2 + (MS - (S^2 - S)/2)8K^2 + 10SK^2 +$ $28K^3/3 - 3K^2 + 8MSK + 8SK$ |
| ZFSL-MPAS | $56K^3/3 - 6K^2 + 8M^2K + 8MK^2 +$ $14SK + 16MSK - 4S^2K + 8SK^2$ |
| MRSL-MPAS | $(MS - (S^2 - S)/2)8K + 10SK +$ $28K^3/3 - 3K^2 + 8SK^2 + 8MSK$ |

Table 6.5: Number of flops required to compute the message vector in massive MIMO (MR precoding) by the proposed antenna selection algorithms highlighted in boldface in Table 6.2; $M$ is the number of BS antennas, $S$ is the number of selected antennas and $K$ is the number of terminals

| **AS Algorithm** | **Complexity (in flops)** |
|---|---|
| ZF-MPAS | $28K^3/3 + 8M^2K + 6SK^2 + 8MSK^2 -$ $4S^2K^2 - 3K^2 + 8MSK + 8SK$ |
| MR-MPAS | $8MK^2 + (MS - (S^2 - S)/2)8K^2 +$ $2SK^2 + 8MSK + 8SK$ |
| ZFSL-MPAS | $28K^3/3 - 3K^2 + 8M^2K + 8MK^2 +$ $14SK + 16MSK - 4S^2K$ |
| MRSL-MPAS | $(MS - (S^2 - S)/2)8K + 10SK + 8MSK$ $(MS - (S^2 - S)/2)8K + 10SK + 8MSK$ |

## 6.3 Concluding Remarks

In this chapter, the proposed algorithms based on the matching pursuit technique were evaluated and also compared with their counterparts via bit error rate. The AS algorithms were tested in different scenarios, comprising rich and poor scattering environments, BPSK and 4-QAM modulations, ZF and MR precoders, and perfect and imperfect CSI knowledge at the BS.

The best AS algorithm choice for massive MIMO systems depends on designer's goal. If one is looking for the best BER performance but at lower power consump-

tion, for instance by activating only a subset of antennas, a symbol-level precoder such as ZFSL-MPAS is a good candidate. On the other hand, if one accepts a certain level of performance but is more concerned with the computational cost, the ZF-MPAS algorithm is a good choice. In addition, if one is planning to use a simple precoding scheme as MR-MPAS, ZF-MPAS algorithm is also a relevant possibility.

This chapter also evaluated the computational complexity of the proposed antenna selection algorithms. The complexity was quantified by the number of flops required to compute the selection vector. Since the computation of the transmitted message is the main task the BS performs in the downlink of a massive MIMO system, we also provided the number of flops needed to compute the transmitted vector. Moreover, the time spent to compute the transmit vector was used to compare the proposed AS algorithms and their counterparts. Indeed, we could observe that the proposed algorithm ZFSL-MPAS is faster than LASSO and MCAS algorithms, in the setup tested. In addition, the proposed ZF-MPAS is also faster than MCAS Algorithm.

# Chapter 7

# Conclusions

## 7.1 Final Remarks

In this work, we explored the antenna selection problem in massive MIMO systems. The criterion employed in massive MIMO systems to select the most promising antennas is based on the maximization of the downlink capacity or on the sparse recovery of a vector of interest. Two algorithms based on the matching pursuit technique are proposed for selecting the subset of active antennas in a massive MIMO system. The algorithms are called Zero Forcing - Matching Pursuit Antenna selection (ZF-MPAS) and Zero Forcing Symbol Level - Matching Pursuit Antenna selection (ZFSL-MPAS). In fact, we also derived two other algorithms based on matching pursuit but idealized for MR precoding. Unfortunately, these MPAS algorithms did not present good results.

The simulation results indicated that ZFSL-MPAS and ZF-MPAS achieve performance level very similar to the ones achieved by their counterparts LASSO and MCAS algorithms, respectively, at lower complexity. As discussed in chapter 6, choosing the best AS algorithm for a massive MIMO system is a task that mainly depends on the level of performance required. For instance, if high BER performance is the most important aspect, we recommend the ZFSL-MPAS algorithm. In contrast, if the computational cost is the main concern, we suggest the ZF-MPAS algorithm that works quite well using both ZF and MR precoding.

## 7.2 Future Work

Massive MIMO is a technique with great promises, but with many challenges and practical issues that researchers still need to investigate. For example, the effect of pilot contamination that seriously limit the performance could be accounted in future simulations.

There are technical aspects that should be considered in order to make massive MIMO a reality. For instance, including constraints regarding the saturation in the power amplifier, low resolution (e.g., one bit) digital-to-analog (DAC) [84, 103]. A possible way is to adapt the Matching Pursuits with Generalized Bit-Planes (MPGBP) algorithm proposed in [23] to select the antennas. Moreover, the MPGBP algorithm can be quite useful to include power allocation in the antenna selection problem. In this way, the elements of the selection vector $\mathbf{z}$ would be allowed to be in the interval [0,1] and hence represent percentage of power. Furthermore, attempts that jointly explore the big three main technologies for 5G: ultra-densification, mmWave (milimeter wave), and massive MIMO. For example, antenna selection and mmWave are worth to be jointly considered [104, 105]. Another relevant issue is the calibration drawback, which is needed in order to compensate for the lack of reciprocity induced by RF components [106].

# Chapter 8

# Antenna Selection in Single-User MIMO

In this chapter we provide a connection between the single-user MIMO system, described in chapter 2, and a centralized sensor network. Such a connection can be useful to see that our low complexity proposed methods can also be employed in sensor networks. Therefore, in this chapter we describe the antenna selection problem in an SU MIMO system, as well as the main algorithms used to select the antennas. This chapter also contain simulation results via BER, MSE and determinant. We also propose a lower complexity algorithm and provide its complexity in number of flops.

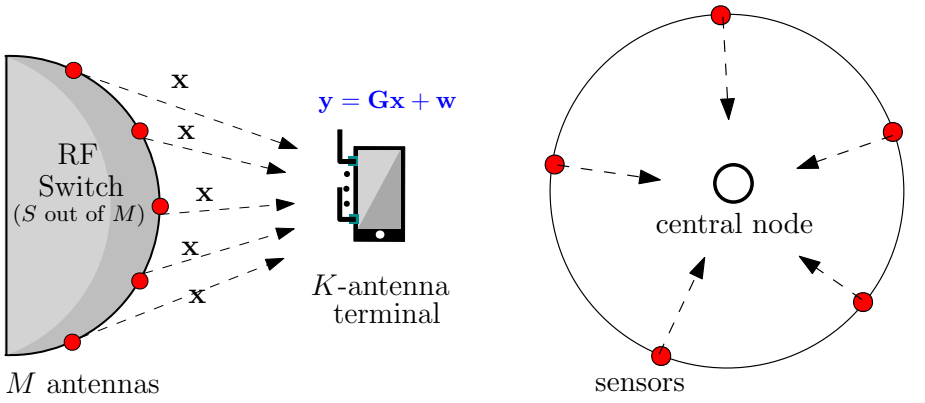
## 8.1 Single-user MIMO versus Sensor selection

Wireless sensor network (WSN) is another system of huge interest in recent research due to applications as target detection, target estimation and target tracking techniques. In a WSN, the sensor nodes transmit their measurements to the central unit, which is responsible for estimating a desired vector by using the sensor measurements received. The $M \times 1$ vector of measurements, received in the central unit, is of the form

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{8.1}$$

where $\mathbf{H}$ is the regressor matrix which represents the effect of the environment in the desired vector $\mathbf{x}$ and $\mathbf{w}$ is a realization of an AWGN zero mean random vector. Since the sensor's resources are limited, sensor selection is an effective way to save power consumption. The problem of selecting the BS antennas in the downlink of a single-user MIMO can be modeled as a centralized sensor selection problem [88–91]. In this way, the sensor nodes play the role of the BS antennas, the central unit is the single-user terminal and the regressor matrix is equivalent to the channel matrix

$\mathbf{H} = \mathbf{G}$, as illustrated in Figure 8.1. The central unit in the WSN must know the regressors in matrix $\mathbf{H}$ in order to estimate vector $\mathbf{x}$. Therefore, the selection criteria used in sensor selection problem usually relies on minimizing functions of the error covariance matrix. As mentioned in subsection 2.2.1, the terminals do not perform channel estimation in massive MIMO due to the precoding and the large number of antennas. However, to use the same selection criteria employed in sensor selection problem, the terminal must have CSI. Thus, exceptionally in this chapter we consider the single-user MIMO system presented in subsection 2.2.1, where precoding is not used.




(a) SU MIMO (Downlink). (b) Centralized sensor network.

Figure 8.1: Comparison between SU MIMO and centralized sensor network schemes.

### 8.1.1 Problem Description

When all the BS antennas are active, the measurements from each antenna are of the form

$$y_m = \mathbf{g}_m^{\mathrm{T}}\mathbf{x} + w_m, \quad m = 1, \ldots M, \tag{8.2}$$

where $\mathbf{g}_m^{\mathrm{T}}$ is the $m$th row of channel matrix $\mathbf{G}$ described in equation (3.1). The transmitted message is represented by vector $\mathbf{x} \in \mathbb{C}^{K\times 1}$. An AWGN, $w_m$, corrupts the message. The element $g_{mk}$ of vector $\mathbf{g}_m$ comprises the fading channel gain between the $m$th BS antenna and the $k$th receive antenna of the terminal. The vector $\mathbf{g}_m$ is considered independent of $\mathbf{x}$ and $w_m$. In vector form,

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w} \tag{8.3}$$

is an $M \times 1$ vector containing the measurements from all active BS antenna elements. For the sake of simplicity, the correlation between the receive antennas is not considered. Hence, the channel matrix $\mathbf{G}$ is composed by vectors $\mathbf{g}_m$ whose entries are generated from a complex Gaussian distribution due to central-limit Theorem [62].

The BS performs the antenna selection as depicted in Figure 8.2, so that the single-user terminal receiver collects only the selected measures,

$$\mathbf{y} = \text{diag}\,(\mathbf{z})\,[\mathbf{G}\widehat{\mathbf{x}} + \mathbf{w}], \tag{8.4}$$

where the selection vector $\mathbf{z}$ is defined in equation (4.1). By pre-multiplying both sides of equation (8.4) by $\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})$ and isolating $\mathbf{x}$, we obtain the estimated transmitted vector

$$\widehat{\mathbf{x}} = (\mathbf{G}^{\text{H}}(\text{diag}\,(\mathbf{z}))^2\mathbf{G})^{-1}(\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{y} - \mathbf{G}^{\text{H}}(\text{diag}\,(\mathbf{z}))^2\mathbf{w}). \tag{8.5}$$

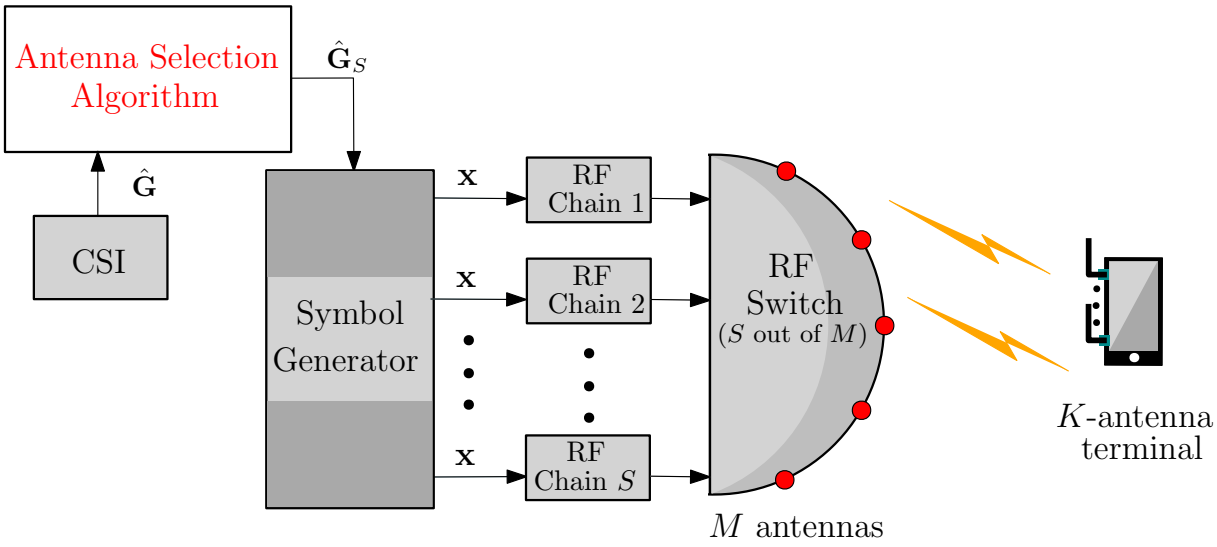



Figure 8.2: Antenna selection process in the downlink of SU MIMO system.

The noise variance is considered known and $(\text{diag}\,(\mathbf{z}))^2 = \text{diag}\,(\mathbf{z})$ because $\mathbf{z}$ is a binary vector. The original transmitted vector can be expressed as

$$\mathbf{x} = (\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G})^{-1}\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{y} \tag{8.6}$$

by zeroing $\mathbf{w}$ in equation (8.5). Thus we can compute the estimation error

$$\mathbf{x} - \hat{\mathbf{x}} = (\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G})^{-1}\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{w}. \tag{8.7}$$

As the noise has zero mean, the estimation error also has zero mean and its covariance matrix is given by

$$\begin{aligned}\boldsymbol{\Sigma} &= (\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G})^{-1}\mathbf{G}^{\text{H}}(\text{diag}\,(\mathbf{z}))^2\sigma^2\mathbf{I}_{(M)}\mathbf{G}(\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G})^{-1} \\ &= \sigma^2(\mathbf{G}^{\text{H}}\text{diag}\,(\mathbf{z})\,\mathbf{G})^{-1} = \left(\frac{1}{\sigma^2}\sum_{m=1}^{M} z_m\mathbf{g}_m\mathbf{g}_m^{\text{H}}\right)^{-1}\end{aligned} \tag{8.8}$$

where $E[\boldsymbol{w}\boldsymbol{w}^T] = \sigma^2\mathbf{I}_{(M)}$ is the covariance matrix of the AWGN vector $\boldsymbol{w}$.

### 8.1.2 Selection Criteria

The selection vector $\mathbf{z}$ is obtained by a criterion that is employed by the antenna selection algorithms. The capacity maximization can also be applied in the SU MIMO context, so that the antennas that most contribute to the maximization in equation (2.3) are chosen. By adding the selection step in equation (2.3), we obtain

$$C_{\mathrm{dl}} = \log\det\left(\mathbf{I}_{(M)} + \frac{\rho_{\mathrm{dl}}}{M}\mathrm{diag}\left(\mathbf{z}\right)\mathbf{G}\mathbf{G}^{\mathrm{H}}\mathrm{diag}\left(\mathbf{z}\right)\right). \tag{8.9}$$

and we can apply the Sylvester theorem, previously used in subsection 4.1.2.1, thus yielding the final expression for the capacity maximization

$$C_{\mathrm{dl}} = \log\det\left(\mathbf{I}_{(K)} + \frac{\rho_{\mathrm{dl}}}{M}\mathbf{G}^{\mathrm{H}}\mathrm{diag}\left(\mathbf{z}\right)\mathbf{G}\right). \tag{8.10}$$

Alternatively, the subset of antennas leading to the minimum estimation error can be selected. To minimize the estimation error, one could try to minimize its variance. However, the estimation error is a vector and its variance is in fact a matrix, well known as covariance matrix. Thus, we aim to minimize an adequate function of the covariance matrix. Fortunately, the theory of optimal experimental design is concerned with the problem of selecting a design which minimizes some function $\psi(\cdot)$ of a matrix $\mathbf{H}$ over all possible designs [92]. The function $\psi(\cdot)$ is called an *optimality criterion* and is usually applied to the covariance matrix $\boldsymbol{\Sigma}$. In this way, the optimality criteria are used to measure the quality of the estimation, where the most common are known as A-and D-optimalities [90]. Both A-and D-optimality criteria are related to the confidence ellipsoid of the estimation error vector $\mathbf{x} - \hat{\mathbf{x}}$

$$\mathcal{E} = (\mathbf{x} - \hat{\mathbf{x}})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \hat{\mathbf{x}}) \tag{8.11}$$

which corresponds to the minimum volume ellipsoid that contains the estimation error vector. In the A-optimality, minimizing the $\mathrm{tr}(\boldsymbol{\Sigma})$ is the same as minimizing the mean-squared-error (MSE) which geometrically corresponds to minimize the sum of the length of each semi-axes of such an ellipsoid. On the other hand, in the D-optimality, maximizing $\det(\boldsymbol{\Sigma}^{-1})$ is equivalent to minimize the log-volume of the confidence ellipsoid [88],

$$\mathrm{logvol}(\mathcal{E}) = \beta - \left(\frac{1}{2}\right)\log\,\det\left(\sum_{m=1}^{M} z_m\mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}\right). \tag{8.12}$$

Observe that maximizing the last term in (8.12) is the same as minimizing the whole

expression. Now we are able to formulate the antenna selection as optimization problems based on A-and D-optimalities.

- A-optimality

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} && \text{tr}\left[\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\text{H}}\right)^{-1}\right] \\ &\text{subject to} && \mathbf{1}^T\mathbf{z} = S \quad z_m \in \{0,1\}, m \in \{1,\cdots,M\} \end{aligned} \tag{8.13}$$

- D-optimality

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} && \log\det\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\text{H}}\right) \\ &\text{subject to} && \mathbf{1}^T\mathbf{z} = S \quad z_m \in \{0,1\}, m \in \{1,\cdots,M\} \end{aligned} \tag{8.14}$$

## 8.2 Convex Optimization Methods for Antenna Selection

### 8.2.1 SU MIMO A-Optimality Convex Problem

The objective function in problem (8.13) is convex, but the binary constraints $z_m \in \{0,1\}$ are not. To obtain a convex optimization problem, we need to replace the binary constraints with inequality constraints. This strategy yields

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{minimize}} && \text{tr}\left[\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\text{H}}\right)^{-1}\right] \\ &\text{subject to} && \mathbf{1}^{\text{T}}\mathbf{z} = S \quad 0 \le z_m \le 1, m \in \{1,\cdots,M\} \end{aligned} \tag{8.15}$$

which is indeed a convex optimization problem. In fact, the problem in equation (8.15) can be transformed to a semidefinite program (SDP), and as a result it is efficiently solved. In SDP one minimizes a linear function subject to the constraint that an affine combination of symmetric matrices is positive semidefinite [50]. We just need to find the linear function to be used as objective function and the

matrices to be used as constraints. We can write the problem (8.15) as

$$\begin{aligned} &\underset{\substack{\mathbf{z}\in\mathbb{R}^{M\times 1}\\ \mathbf{u}\in\mathbb{R}^{S\times 1}}}{\text{minimize}} \quad \mathbf{1}^{\mathrm{T}}\mathbf{u} \\ &\text{subject to} \qquad u_s \geq \mathbf{e}_s^{\mathrm{T}} \left( \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \right)^{-1} \mathbf{e}_s \quad s \in \{1, \cdots, S\} \\ &\mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, \quad m \in \{1, \cdots, M\} \end{aligned} \tag{8.16}$$

where the vector $\mathbf{e}_s$ is the $s$th column of the identity matrix $\mathbf{I}_{(S)}$. An efficient way to solve the problem in equation (8.16) is to define a matrix

$$\mathbf{B}_s = \begin{bmatrix} u_s & \mathbf{e}_s^{\mathrm{T}} \\ \mathbf{e}_s & \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \end{bmatrix} \tag{8.17}$$

for $s \in \{1, \cdots, S\}$, where $u_s$ is the $s$th element of an auxiliary vector $\mathbf{u}$. The Schur complement [50] of matrix $\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}$ in $\mathbf{B}_s$ is

$$\chi_j = u_s - \mathbf{e}_s^{\mathrm{T}} \left( \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \right)^{-1} \mathbf{e}_s. \tag{8.18}$$

One of the Schur complement' properties states that if matrix $\mathbf{B}_s$ is semidefinite positive, $\mathbf{B}_s \succeq 0$, then $\chi_s \geq 0$ [50]. Therefore, it is straightforward to write

$$\mathbf{e}_s^{\mathrm{T}} \left( \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \right)^{-1} \mathbf{e}_s \leq u_s \tag{8.19}$$

and as a result the problem in equation 8.15 can be written as

$$\begin{aligned} &\underset{\substack{\mathbf{z}\in\mathbb{R}^{M\times 1}\\ \mathbf{u}\in\mathbb{R}^{S\times 1}}}{\text{minimize}} \quad \mathbf{1}^{\mathrm{T}}\mathbf{u} \\ &\text{subject to} \qquad \begin{bmatrix} u_s & \mathbf{e}_s^{\mathrm{T}} \\ \mathbf{e}_s & \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \end{bmatrix} \succeq 0 \quad s \in \{1, \cdots, S\} \\ &\mathbf{1}^{T}\mathbf{z} = S \quad z_m \geq 0, \quad m \in \{1, \cdots, M\} \end{aligned} \tag{8.20}$$

which is an SDP problem and can be solved by interior-point methods [50]. The resulting antenna selection is detailed in Algorithm 10, where a procedure very similar to the Algorithm 1 is employed.

**Algorithm 10** : A-Optimality Antenna Selection (AOAS) [91]

1) Input: $\mathbf{G}$, $S$
2) Initialization: $\mathcal{I}' = \{1, 2, \dots M\}$, $\mathcal{I} = \{\}$
3) Find $\mathbf{z}$ that solves the problem (equation (8.20))

$$\begin{aligned} &\underset{\substack{\mathbf{z}\in\mathbb{R}^{M\times 1}\\ \mathbf{u}\in\mathbb{R}^{S\times 1}}}{\text{minimize}} \quad \mathbf{1}^{\mathrm{T}}\mathbf{u} \\ &\text{subject to} \quad \begin{bmatrix} u_s & \mathbf{e}_s^{\mathrm{T}} \\ \mathbf{e}_s & \sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} \end{bmatrix} \succeq 0 \quad s \in \{1, \cdots, S\} \\ &\mathbf{1}^T\mathbf{z} = S \quad z_m \geq 0, \quad m \in \{1, \cdots, M\} \end{aligned}$$

4) Store the indices of the $k$-largest $\mathbf{z}$ in $\mathcal{I}$
5) $\mathcal{I}' = \mathcal{I}' - \mathcal{I}$
6) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
7) Set $z_i \leftarrow 0$ for $i \in \mathcal{I}'$
8) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

9) Output: $\mathbf{G}_S$

### 8.2.2 SU MIMO D-Optimality Convex Problem

The objective function of the problem in equation (8.14) is also concave, but the binary constraints are non-convex. As in the A-Optimality case, the binary constraints in equation (8.14) can also be replaced by the inequalities yielding the convex relaxed version of problem (8.14):

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad \log\det\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}\right) \\ &\text{subject to} \quad \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, m \in \{1, \cdots, M\} \end{aligned} \tag{8.21}$$

which can be solved by interior-point methods [50]. Similarly to the massive MIMO maximum capacity case, the log-barrier method [88] can be used to obtain an approximated version of the problem in equation (8.21),

$$\begin{aligned} \underset{\mathbf{z}\in\mathcal{R}^{M\times 1}}{\text{maximize}} \quad & \log\det\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}\right) + \\ & \kappa \sum_{m=1}^{M} (\log(z_m) + \log(1 - z_m)) \\ \text{subject to} \quad & \mathbf{1}^T\mathbf{z} = S \end{aligned} \tag{8.22}$$

where $\kappa$ is a positive constant which controls the quality of the approximation. The optimization problem in (8.22) can also be solved by applying the Newton Method [50]. The resulting antenna selection is summarized in Algorithm 11.

**Algorithm 11** : D-Optimality Antenna Selection (DOAS) [88]

1) Input: $\mathbf{G}$, $S$
2) Initialization: $\mathcal{I}' = \{1, 2, \dots M\}$, $\mathcal{I} = \{\}$
3) Find $\mathbf{z}$ that solves the problem (equation (8.21))

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad & \log\det\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}\right) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad 0 \le z_m \le 1, m \in \{1, \cdots, M\} \end{aligned}$$

4) Store the indices of the $k$-largest $\mathbf{z}$ in $\mathcal{I}$
5) $\mathcal{I}' = \mathcal{I}' - \mathcal{I}$
6) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
7) Set $z_i \leftarrow 0$ for $i \in \mathcal{I}'$
8) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

9) Output: $\mathbf{G}_S$

### 8.2.3 SU MIMO Downlink Capacity Convex Problem

The downlink capacity of an SU MIMO system was previously defined in equation (2.3). Since no user power allocation is performed in SU MIMO, because there is only one user, the problem in equation (2.3) can be seen as a simplified version of the problem in equation (4.8). After performing the relaxation in the binary constraints, we can solve the problem

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\mathrm{H}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G})) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad 0 \le z_m \le 1, m = \{1, \cdots, M\} \end{aligned} \tag{8.23}$$

by using interior-point methods, or we can use an approximation via log-barrier approach which yields the problem

$$\begin{aligned} \underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad & \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\mathrm{H}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}))+ \\ & \kappa\sum_{m=1}^{M}(\log(z_m) + \log(1 - z_m)) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad m \in \{1, \cdots, M\} \end{aligned} \tag{8.24}$$

that can be solved by Newton method [50]. The resulting antenna selection is summarized in Algorithm 12.

**Algorithm 12** : Max Capacity Antenna Selection (MCAS) [20]

1) Input: $\mathbf{G}$, $S$
2) Initialization: $\mathcal{I}' = \{1, 2, \dots M\}$, $\mathcal{I} = \{\}$
3) Find $\mathbf{z}$ that solves the problem (equation (8.23))

$$\begin{aligned} &\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} && \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}^{\mathrm{H}}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G})) \\ &\text{subject to} && \mathbf{1}^{\mathrm{T}}\mathbf{z} = S \quad 0 \leq z_m \leq 1, m = \{1, \cdots, M\} \end{aligned}$$

4) Store the indices of the $k$-largest $\mathbf{z}$ in $\mathcal{I}$
5) $\mathcal{I}' = \mathcal{I}' - \mathcal{I}$
6) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
7) Set $z_i \leftarrow 0$ for $i \in \mathcal{I}'$
8) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

9) Output: $\mathbf{G}_S$

## 8.3 Low Complexity Antenna Selection

Although a good performance is achieved by antenna selection strategies solved via convex optimization, a price is paid in computational cost. Therefore, different sub-optimal antenna selection methods have been proposed recently, as described in [97] and references therein. We present in this section some existing low complexity antenna selection methods. Moreover, we propose a low complexity antenna selection algorithm based on the approximation of the D-optimality function.

### 8.3.1 Trace-Based Low Complexity Problem

In the A-optimal design approach, the objective is to select the antennas so that the trace of the resulting covariance matrix is minimized. A related method is proposed in [97], in which the minimization of $\mathrm{tr}[(\mathbf{G}\mathbf{G}^{\mathrm{H}})^{-1}]$ is also the main goal. The authors in [97] start by considering that the full-rank $M \times K$ matrix $\mathbf{G}$ with QR decomposition $\mathbf{G} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q}$ is an $M \times K$ matrix of orthonormal column vectors, and $\mathbf{R}$ is an upper triangular square matrix. Therefore, the trace of $(\mathbf{G}\mathbf{G}^{\mathrm{H}})^{-1}$ can be expressed as

$$\mathrm{tr}[(\mathbf{G}\mathbf{G}^{\mathrm{H}})^{-1}] = \sum_{k=1}^{K} T_k, \tag{8.25}$$

where

$$T_k = \frac{1 + \sum_{t=1}^{k-1} |p_{k,t}|^2}{|r_{k,k}|}. \tag{8.26}$$

In which $r_{i,j}$ denotes the $(i,j)$th entry of $\mathbf{R}$, and

$$p_{k,t} = \frac{r_{t,k} - \sum_{j=t+1}^{k-1} p_{k,j} r_{t,j}}{r_{t,t}}. \tag{8.27}$$

In the proposed trace-based AS scheme, the antenna that results in the smallest value of $T_k$ is selected at the $k$th step. The process continues until $S$ antennas are selected. Resulting in the Algorithm 13.

---

**Algorithm 13** : Trace-Based Antenna Selection (TBAS) [97]

---

1) $\mathcal{I}' \leftarrow \{1,2,\dots M\}$
2) $\mathcal{I} \leftarrow \{\}$, $\mathbf{z} = \mathbf{0}_M$
3) **for** $i \in \mathcal{I}'$ **do**
$\mathbf{v}_i \leftarrow \mathbf{h}_i$
$r_{1,i} \leftarrow |\mathbf{v}_i|$
$T_i \leftarrow 1/|r_{1,i}|^2$
**end for**
4) **for** $k = 1, \dots, S$ **do**
a)Choose the best antenna
$k_{\text{opt}} \leftarrow \text{argmin}_{i \in \mathcal{I}'} T_i$
$\mathcal{I}' \leftarrow \mathcal{I}' - \{k_{\text{opt}}\}$
$\mathcal{I} \leftarrow \mathcal{I} \cup \{k_{\text{opt}}\}$
b)Perform Gram-Schmidt Orthogonalization
$\mathbf{u} \leftarrow \mathbf{v}_{k_{\text{opt}}}/r_{k,k_{\text{opt}}}$
**for** $i \in \mathcal{I}'$ **do**
$r_{k,i} \leftarrow \mathbf{u}^{\text{H}}\mathbf{v}_i$
$\mathbf{v}_i \leftarrow \mathbf{v}_i - r_{k,i}\mathbf{u}$
$r_{k+1,i} \leftarrow |\mathbf{v}_i|$
**for** $t = k, k-1, \dots, 1$ **do**
$p_{i,t} = (r_{t,i} - \sum_{j=t+1}^{k} p_{i,j} r_{t,j})/r_{t,t}$
**end for**
$T_i = (1 + \sum_{t=1}^{k} |p_{i,t}|^2)/|r_{k+1,i}|^2$
**end for**
**end for**
5) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
6) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \text{rem}(\text{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\text{T}}\mathbf{G}$$

7) Output: $\mathbf{G}_S$

---

### 8.3.2 Determinant-Based Low Complexity Problem

In the D-optimal design, the main goal is to find the vector $\mathbf{z}$ that solves

$$\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad \log\det\left(\sum_{m=1}^{M} z_m \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}\right) \tag{8.28}$$

in which $\sum_{m=1}^{M} \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}} = \mathbf{G}\mathbf{G}^{\mathrm{H}}$ is the inverse of the covariance matrix. Such a problem is similar to the maximization of the massive MIMO downlink capacity,

$$\underset{\mathbf{z}\in\mathbb{R}^{M\times 1}}{\text{maximize}} \quad \log_2(\det(\mathbf{I}_{(K)} + \mathbf{G}\mathrm{diag}\,(\mathbf{z})\,\mathbf{G}^{\mathrm{H}})). \tag{8.29}$$

The matrix determinant lemma [98], which states that

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^{\mathrm{H}}) = (1 + \mathbf{v}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{u})\det(\mathbf{A}), \tag{8.30}$$

can be used to express the determinant of $\mathbf{A}_0 = \sum_{m=1}^{M} \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}$ as

$$\det(\mathbf{A}_0) = \prod_{i=1}^{M} \left[1 + \mathbf{g}_{M-i+1}^{\mathrm{H}} (\mathbf{A}_{i-1})^{-1} \mathbf{g}_{M-i+1}\right] \tag{8.31}$$

where $\mathbf{A}_{i-1} = \sum_{m=1}^{M-i} \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}$. If we try to calculate each product element in equation (8.31), it would lead to the computation of $M$ matrix inversions. We can compute an auxiliary vector and search for the $S$ largest vector components so that the selected antenna indices match the largest elements indices. Hence, we can use an approximation and form the new auxiliary vector,

$$\mathbf{t} = \left[\mathbf{g}_1^{\mathrm{H}} (\mathbf{A}_0)^{-1} \mathbf{g}_1, \ldots, \mathbf{g}_M^{\mathrm{H}} (\mathbf{A}_0)^{-1} \mathbf{g}_M\right]^{\mathrm{T}} \tag{8.32}$$

where we only need to invert the matrix $\mathbf{A}_0 = \sum_{m=1}^{M} \mathbf{g}_m \mathbf{g}_m^{\mathrm{H}}$ once, and thus avoiding the $M$ matrix inversions. The resulting algorithm is described in Table 14. Alternatively, the authors in [86] used the matrix inversion lemma, which is derived from the Sherman Morrison formula to alleviate the computation of the $M$ matrix inversions. The inverse matrix

$$\mathbf{B}_n = \left(\mathbf{I}_K + \mathbf{G}\mathbf{G}^{\mathrm{H}}\right)^{-1} \tag{8.33}$$

is updated whenever a row is added, based on the previous inverse matrix $\mathbf{B}_{n-1}$. Hence resulting in the Algorithm 15.

**Algorithm 14** : Determinant Approximation Antenna Selection (DAAS)

1) Input: $\mathbf{G}$, $S$
2) Initialization: $\mathbf{z} = \mathbf{0}_M$, $\mathcal{I} = \{\}$
3) Compute auxiliary vector

$$\mathbf{t} = \left[\mathbf{g}_1^{\mathrm{H}}(\mathbf{A}_0)^{-1}\mathbf{g}_1, \ldots, \mathbf{g}_M^{\mathrm{H}}(\mathbf{A}_0)^{-1}\mathbf{g}_M\right]^{\mathrm{T}}$$

4) Store the indices of the $k$-largest $\mathbf{t}$ in $\mathcal{I}$
5) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
6) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

7) Output: $\mathbf{G}_S$

---

**Algorithm 15** : Fast Antenna Selection (FAS) [86]

1) Input: $\mathbf{G}$, $S$, $\rho_{\mathrm{dl}}$
2) Initialization: $\mathbf{z} = \mathbf{0}_M$, $\mathcal{I} = \{\}$, $\mathbf{B} = \mathbf{I}_K$
3) **for** $j = 1, \ldots, M$ **do**
$\alpha_j \leftarrow \mathbf{g}_j^{\mathrm{H}}\mathbf{g}_j$
**end for**
4) **for** $n = 1, \ldots, S$ **do**
$J \leftarrow \ \mathrm{argmax}_{j \in \mathcal{I}} \alpha_j$
$\mathcal{I} \leftarrow \mathcal{I} \cup J$
**if** $n < S$
$\mathbf{a} \leftarrow \frac{1}{\sqrt{M/\rho_{\mathrm{dl}} + \alpha_j}} \mathbf{B}\mathbf{g}_j$
$\mathbf{B} \leftarrow \mathbf{B} - \mathbf{a}\mathbf{a}^{\mathrm{H}}$
for all $j \in \mathcal{I}$
$\alpha_j \leftarrow \alpha_j - |\mathbf{a}^{\mathrm{H}}\mathbf{g}_j|^2$
end for
**end if**
**end for**
5) Set $z_i \leftarrow 1$ for $i \in \mathcal{I}$
6) Compute the $S$-selected channel matrix

$$\mathbf{G}_S = \mathrm{rem}(\mathrm{diag}\,(\mathbf{z}))\mathbf{G} = \mathbf{Z}^{\mathrm{T}}\mathbf{G}$$

7) Output: $\mathbf{G}_S$

---

## 8.4 Single-user MIMO simulations

In this section, the performances of the AS algorithms are compared in the downlink of a single-user MIMO system. The system is composed of a unique terminal equipped with $K = 20$ antennas which is served by a BS equipped with $M = 100$ antennas in a single-cell. Analogously, the system can be seen as a central unit that estimates a message of size $K$ from $M$ estimations received by $M$ sensor

nodes in the network.

The BS transmits a block of 200 message vectors $[\mathbf{x}_1 \quad \dots \quad \mathbf{x}_{200}]$ to the single terminal in the cell. Each vector $\mathbf{x}_i \in \mathbb{R}^{K\times 1}$ in the block is composed by BPSK symbols. The terminal is assumed to know CSI perfectly and then is able to estimate each $\mathbf{x}_i$ in the block by computing

$$\mathbf{x}_i = (\mathbf{G}_S^{\mathrm{H}}\mathbf{G}_S)^{-1}\mathbf{G}_S^{\mathrm{H}}\mathbf{y}_i \tag{8.34}$$

where $\mathbf{G}_S$ is the $S$ selected channel matrix and $\mathbf{y}_i, i \in \{1, \cdots, 200\}$ is the received vector. The matrix $\mathbf{G}_S$ is generated by the AS algorithms considered in the simulations, which are listed in Table 8.1. The proposed algorithms are highlighted in boldface in Table 8.1. Observe that we are also considering the antenna selection algorithms based on matching pursuit.

Since the different scenarios were already tested in massive MIMO system, here in SU MIMO system, we only consider scenario 9 which is depicted in Figures 8.3 and 8.4. Such a basic scenario is enough to illustrate the single-user MIMO behavior and that can be propagate to sensor networks. In scenario 9, the AS algorithms are compared in three different ways: BER in Figure 8.3a, A-optimality (MSE) in Figure 8.4a and D-optimality (determinant) in Figure 8.4c. To better compare the algorithms we also provide a zoom in Figures 8.3b, 8.4b and 8.4d.

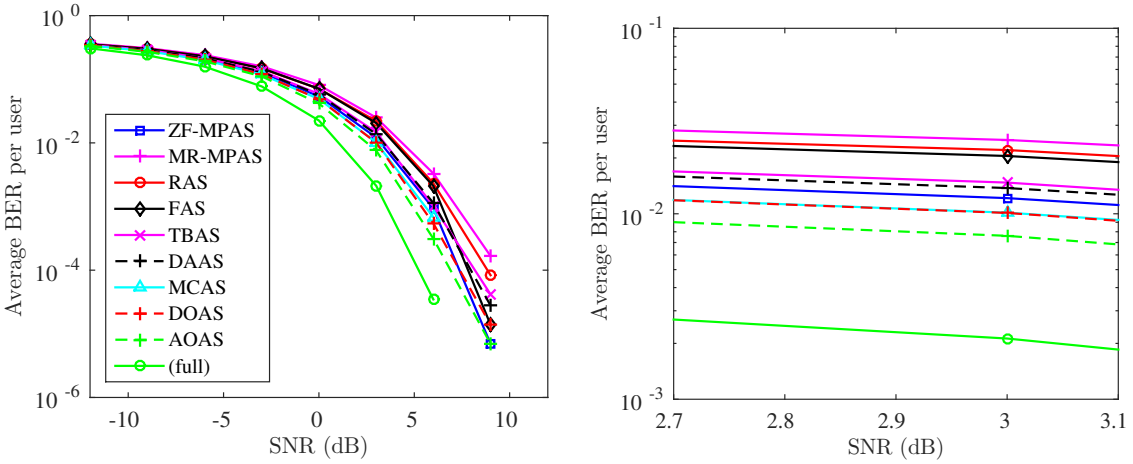
As expected, the methods based on convex optimization AOAS, DOAS and MCAS are the ones that achieved the best performances in general. For example, AOAS is the one that achieved the lowest MSE level in Figures 8.4a and 8.4b and also has the best BER performance, which is reasonable since AOAS employs the trace as objective function to be minimized. DOAS and MCAS algorithms yielded the best performance in terms of the determinant, which again is justifiable as they both maximize the determinant in their optimization problems. Nevertheless, these methods are too expensive in terms of computation. Among the low complexity AS algorithms, the proposed ZF-MPAS is the one that achieved the best BER performance. In terms of MSE and determinant, the proposed DAAS is the best option in the group of low complexity AS algorithms.

### 8.4.1 Computational complexity

In this subsection we compute the complexity in flops needed by the proposed DAAS algorithm. To compute the auxiliary vector detailed in equation (8.32) we need to calculate $2M$ matrix-vector products. First of all, we need to obtain the square matrix $\left(\sum_{m=1}^{M} \mathbf{g}_m\mathbf{g}_m^{\mathrm{H}}\right)^{-1}$ of size $K$, which requires $28K^3/3 - 3K^2$ flops, according to subsection 6.2.1.5. Fortunately we need to compute this inverse only once. Now we are able to compute the matrix-vector products, which involves $16MK^2$ flops

Table 8.1: Antenna selection algorithms evaluated in the SU-MIMO simulations

| **AS Algorithm** | **Description** | **Location** |
|---|---|---|
| **ZF-MPAS** | Zero forcing - Matching Pursuits Antenna Selection | Algorithm 7 |
| **MR-MPAS** | Maximum Ratio - Matching Pursuits Antenna Selection | Algorithm 6 |
| AOAS [91] | A-Optimality Antenna Selection | Algorithm 10 |
| DOAS [88] | D-Optimality Antenna Selection | Algorithm 11 |
| TBAS [97] | Trace Based Antenna Selection | Algorithm 13 |
| **DAAS** | Determinant Approximation Antenna Selection | Algorithm 14 |
| FAS [86] | Fast Antenna Selection | Algorithm 15 |
| RAS | Random Antenna Selection | Algorithm 3 |
| MCAS [20] | Maximum Capacity Antenna Selection | Algorithm 12 |



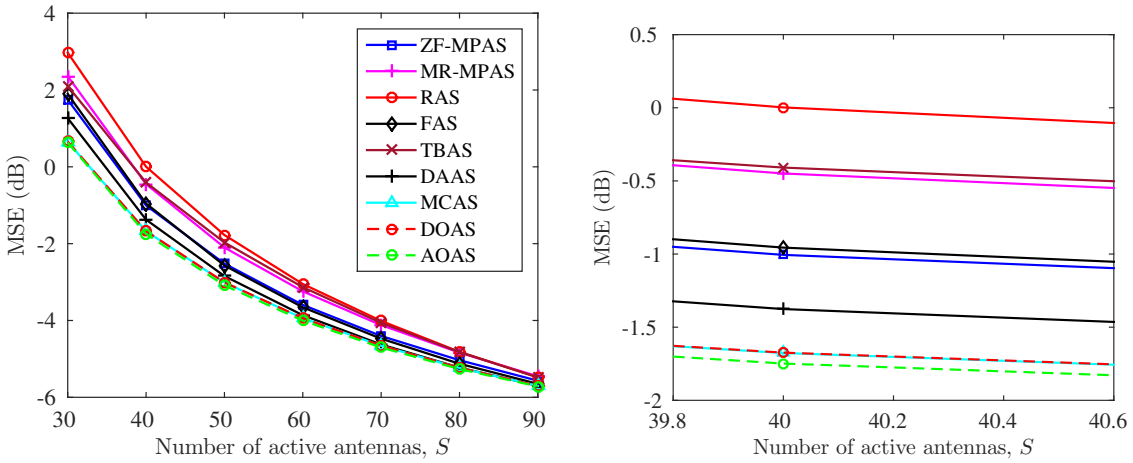

(a) BER, $S = 34$.

(b) BER (zoom), $S = 34$.

Figure 8.3: Scenario 9: Single-user MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols.

in total. Then, the DAAS algorithm requires $28K^3/3 - K^2(3 - 16M)$ flops to build the selection vector $\mathbf{z}$.

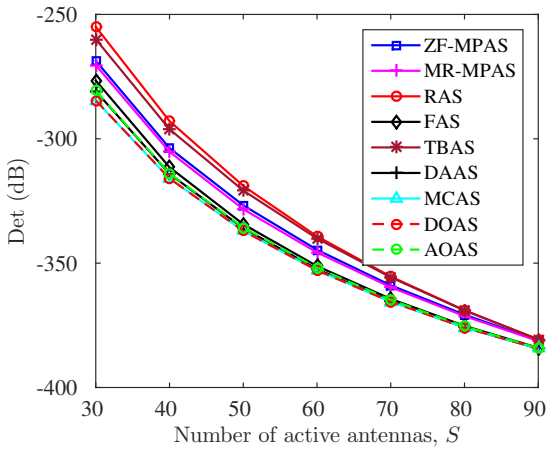
Figure 8.5 illustrates the time spent by each AS algorithm to produce the selection vector $\mathbf{z}$ in the downlink of a single-user MIMO system. Indeed, DOAS and MCAS algorithms, which are based on convex optimization, are more computationally intensive as shown in Figure 8.5. The least computationally intensive algorithms are the ones proposed in this work as can be seen in Figure 8.5.



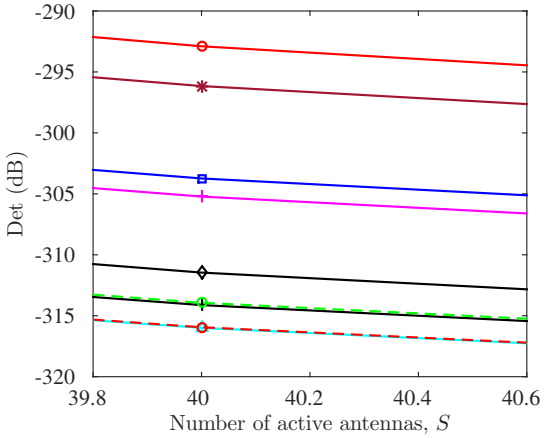

(a) MSE




(b) MSE (zoom).




(c) Determinant.




(d) Determinant (zoom).

Figure 8.4: Scenario 9: Single-user MIMO system with perfect CSI knowledge, uncorrelated Rayleigh channel, and BPSK transmitted symbols, SNR = 10 dB.

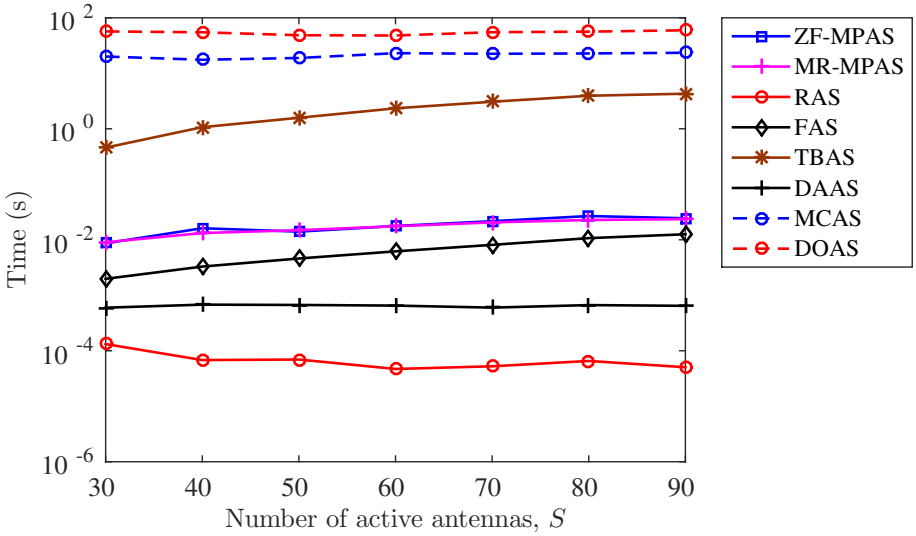



Figure 8.5: Time spent to compute the selection vector by each AS algorithm in an SU MIMO system.

## 8.5 Concluding Remarks

In this chapter, we observed how an SU MIMO system is related to a centralized sensor network. We briefly summarized the existing antenna selection algorithms based on convex optimization and low complexity approaches for an SU MIMO system. In both SU MIMO and sensor networks we can use a criterion based on the minimization of the estimation error at the terminal. Although there is no theory motivation to employ this latter criterion in massive MIMO, we tried it. However, no success was obtained and hence they are not shown in the simulation results for Massive MIMO. We also employed the algorithms originally conceived for massive MIMO in the SU MIMO simulations, and in this setup, we obtained interesting results. We also develop one algorithm based on approximation of a function of the error covariance matrix, which we called DAAS algorithm. We consider DAAS the best option among the low complexity AS algorithms as it performs well enough in the three different aspects considered in the simulations. Moreover, in the sensor network context, DAAS is also the best candidate since its complexity is low without losing too much in performance.

# Bibliography


[1] CISCO. "Visual networking index", *white paper at Cisco.com*, March 2017. Available at `https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html`. [Online; accessed 9-August-2018].

[2] Available at `https://en.wikipedia.org/wiki/Mobile_radio_telephone#/media/File:Mobile_radio_telephone.jpg`. [Online; accessed 9-August-2018].

[3] Available at `https://images-na.ssl-images-amazon.com/images/I/61g7lqLaxWL._SL1200_.jpg`. [Online; accessed 9-August-2018].

[4] KORHONEN, J. *Introduction to 4G mobile communications*. Artech House, 2014.

[5] ALSHARIF, M. H., NORDIN, R. "Evolution towards fifth generation (5G) wireless networks: Current trends and challenges in the deployment of millimetre wave, massive MIMO, and small cells", *Telecommunication Systems*, v. 64, n. 4, pp. 617–637, 2017.

[6] ZHANG, Z., WANG, X., LONG, K., et al. "Large-scale MIMO-based wireless backhaul in 5G networks", *IEEE Wireless Communications*, v. 22, n. 5, pp. 58–66, October 2015.

[7] BOGALE, T. E., LE, L. B. "Massive MIMO and mmWave for 5G Wireless HetNet: Potential Benefits and Challenges", *IEEE Vehicular Technology Magazine*, v. 11, n. 1, pp. 64–75, March 2016.

[8] CHEN, S., ZHAO, J. "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication", *IEEE Communications Magazine*, v. 52, n. 5, pp. 36–43, May 2014.

[9] CHIN, W. H., FAN, Z., HAINES, R. "Emerging Technologies and Research Challenges for 5G Wireless Networks", *IEEE Wireless Communications*, v. 21, pp. 106–112, 04 2014.

[10] BANGERTER, B., TALWAR, S., AREFI, R., et al. "Networks and devices for the 5G era", *IEEE Communications Magazine*, v. 52, n. 2, pp. 90–96, February 2014.

[11] JUNGNICKEL, V., MANOLAKIS, K., ZIRWAS, W., et al. "The role of small cells, coordinated multipoint, and massive MIMO in 5G", *IEEE Communications Magazine*, v. 52, n. 5, pp. 44–51, May 2014.

[12] CORSON, M. S., LAROIA, R., LI, J., et al. "Toward proximity-aware internetworking", *IEEE Wireless Communications*, v. 17, n. 6, pp. 26–33, 2010.

[13] MAEDER, A., ROST, P., STAEHLE, D. "The challenge of M2M communications for the cellular radio access network". In: *Proc. Würzburg Workshop IP, Joint ITG Euro-NF Workshop "Vis. Future Gener. Netw." EuroView*, pp. 1–2, 2011.

[14] ANDREWS, J. G., BUZZI, S., CHOI, W., et al. "What will 5G be?" *IEEE Journal on selected areas in communications*, v. 32, n. 6, pp. 1065–1082, 2014.

[15] ANDERSON, C. R., RAPPAPORT, T. S. "In-building wideband partition loss measurements at 2.5 and 60 GHz", *IEEE transactions on wireless communications*, v. 3, n. 3, pp. 922–928, 2004.

[16] ALEJOS, A. V., SANCHEZ, M. G., CUINAS, I. "Measurement and analysis of propagation mechanisms at 40 GHz: Viability of site shielding forced by obstacles", *IEEE Transactions on Vehicular Technology*, v. 57, n. 6, pp. 3369–3380, 2008.

[17] LARSSON, E. G., EDFORS, O., TUFVESSON, F., et al. "Massive MIMO for next generation wireless systems", *IEEE communications magazine*, v. 52, n. 2, pp. 186–195, February 2014.

[18] MARZETTA, T. L., LARSSON, E. G., YANG, H., et al. *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[19] FELLER, W. "Law of large numbers for identically distributed variables", *An introduction to probability theory and its applications*, v. 2, pp. 231–234, 1971.

[20] GAO, X., EDFORS, O., LIU, J., et al. "Antenna selection in measured massive MIMO channels using convex optimization", *Globecom Workshops (GC Wkshps), 2013 IEEE*, pp. 129–134, December 2013.

[21] CHAVES, R. S. "Joint Precoding And Antenna Selection In Massive MIMO Systems", *M.Sc. dissertation*, March 2018. COPPE/UFRJ.

[22] MALLAT, S. *A wavelet tour of signal processing: the sparse way*. Elsevier/Academic Press, 2009.

[23] DA SILVA, E. A. B., LOVISOLO, L., DUTRA, A. J., et al. "FIR filter design based on successive approximation of vectors", *IEEE Transactions on Signal Processing*, v. 62, n. 15, pp. 3833–3848, May 2014.

[24] HUANG, H., PAPADIAS, C. B., VENKATESAN, S. *MIMO Communication for cellular Networks*. Springer Science & Business Media, 2011.

[25] RING, D. H. "Mobile Telephony - Wide Area Coverage", *Bell Laboratories Technical Memorandum*, December 1947.

[26] BJÖRNSON, E., HOYDIS, J., SANGUINETTI, L. "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency", *Foundations and Trends® in Signal Processing*, v. 11, n. 3-4, pp. 154–655, November 2017.

[27] SWINDLEHURST, A. L., AYANOGLU, E., HEYDARI, P., et al. "Millimeter-wave massive MIMO: The next wireless revolution?" *IEEE Communications Magazine*, v. 52, n. 9, pp. 56–62, September 2014.

[28] WALKE, B. H., MANGOLD, S., BERLEMANN, L. *IEEE 802 wireless systems: protocols, multi-hop mesh/relaying, performance and spectrum coexistence*. John Wiley & Sons, 2007.

[29] CHAN, P. W., LO, E. S., WANG, R. R., et al. "The evolution path of 4G networks: FDD or TDD?" *IEEE Communications Magazine*, v. 44, n. 12, pp. 42–50, December 2006.

[30] ALEXANDERSON, E. F. "Transoceanic radio communication", *Presented at a joint meeting of the American Institute of Electrical Engineers and the Institute of Radio Engineers*, New York, NY, October 1919.

[31] KAYE, A., GEORGE, D. "Transmission of multiplexed PAM signals over multiple channel and diversity systems", *IEEE Transactions on Communication Technology*, v. 18, n. 5, pp. 520–526, October 1970.

[32] SALZ, J. "Digital transmission over cross-coupled linear channels", *AT&T Technical Journal*, v. 64, n. 6, pp. 1147–1159, 1985.

[33] TAROKH, V., SESHADRI, N., CALDERBANK, A. R. "Space-time codes for high data rate wireless communication: Performance criterion and code construction", *IEEE transactions on information theory*, v. 44, n. 2, pp. 744–765, March 1998.

[34] ALAMOUTI, S. M. "A simple transmit diversity technique for wireless communications", *IEEE Journal on selected areas in communications*, v. 16, n. 8, pp. 1451–1458, October 1998.

[35] RALEIGH, G. G., CIOFFI, J. M. "Spatio-temporal coding for wireless communication", *IEEE Transactions on communications*, v. 46, n. 3, pp. 357–366, March 1998.

[36] FOSCHINI, G. J., GANS, M. J. "On limits of wireless communications in a fading environment when using multiple antennas", *Wireless personal communications*, v. 6, n. 3, pp. 311–335, 1998.

[37] MARZETTA, T. L., HOCHWALD, B. M. "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading", *IEEE transactions on Information Theory*, v. 45, n. 1, pp. 139–157, January 1999.

[38] GUEY, J.-C., FITZ, M. P., BELL, M. R., et al. "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels", *IEEE Transactions on Communications*, v. 47, n. 4, pp. 527–537, April 1999.

[39] TAROKH, V., JAFARKHANI, H., CALDERBANK, A. R. "Space-time block codes from orthogonal designs", *IEEE Transactions on information theory*, v. 45, n. 5, pp. 1456–1467, July 1999.

[40] TELATAR, E. "Capacity of Multi-antenna Gaussian Channels", *Transactions on Emerging Telecommunications Technologies*, v. 10, n. 6, pp. 585–595, 1999.

[41] RUSEK, F., PERSSON, D., LAU, B. K., et al. "Scaling up MIMO: Opportunities and challenges with very large arrays", *IEEE Signal Processing Magazine*, v. 30, n. 1, pp. 40–60, January 2013.

[42] GORSHE, S., RAGHAVAN, A., GALLI, S., et al. "Broadband access: wireline and wireless-alternatives for internet services", 2014.

[43] CHIANI, M., WIN, M. Z., ZANELLA, A. "On the capacity of spatially correlated MIMO Rayleigh-fading channels", *IEEE Transactions on Information Theory*, v. 49, n. 10, pp. 2363–2371, October 2003.

[44] FOSCHINI, G. J. "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas", *Bell Labs Technical Journal*, v. 1, n. 2, pp. 41–59, August 1996.

[45] ZHANG, X., MOLISCH, A. F., KUNG, S. Y. "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection", *IEEE Transactions on Signal Processing*, v. 53, n. 11, pp. 4091–4103, November 2005.

[46] VENKATESWARAN, V., VAN DER VEEN, A. J. "Analog Beamforming in MIMO Communications With Phase Shift Networks and Online Channel Estimation", *IEEE Transactions on Signal Processing*, v. 58, n. 8, pp. 4131–4143, August 2010.

[47] GESBERT, D., KOUNTOURIS, M., HEATH, R. W., et al. "Shifting the MIMO Paradigm", *IEEE Signal Processing Magazine*, v. 24, n. 5, pp. 36–46, September 2007.

[48] KNOPP, R., HUMBLET, P. "Information capacity and power control in single-cell multiuser communications". In: *Proceedings IEEE International Conference on Communications ICC*, v. 1, pp. 331–335, Seattle, WA, USA, USA, June 1995. IEEE.

[49] BROWN, T., KYRITSI, P., DE CARVALHO, E. *Practical guide to MIMO radio channel: With MATLAB examples*. John Wiley & Sons, 2012.

[50] BOYD, S., VANDENBERGHE, L. *Convex Optimization*. Cambridge university press, 2004.

[51] DINIZ, P. S. R., MARTINS, W. A., LIMA, M. V. S. "Block Transceivers: OFDM and Beyond", *Synthesis Lectures on Communications*, v. 5, n. 1, pp. 1–206, June 2012.

[52] MARZETTA, T. L. "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas", *IEEE Transactions on Wireless Communications*, v. 9, n. 11, pp. 3590–3600, November 2010.

[53] LU, L., LI, G. Y., SWINDLEHURST, A. L., et al. "An Overview of Massive MIMO: Benefits and Challenges", *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, pp. 742–758, October 2014.

[54] TANG, M. F., SU, B. "Downlink Precoding for Multiple Users in FDD Massive MIMO Without CSI Feedback", *Journal of Signal Processing Systems*, v. 83, n. 2, pp. 151–163, May 2016.

[55] NAM, J., AHN, J. Y., ADHIKARY, A., et al. "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information". In: *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, Princeton, NJ, USA, March 2012. IEEE.

[56] RAO, X., LAU, V. K. N. "Distributed Compressive CSIT Estimation and Feedback for FDD Multi-user Massive MIMO Systems", *IEEE Transactions on Signal Processing*, v. 62, n. 12, pp. 3261–3271, May 2014.

[57] BJÖRNSON, E., LARSSON, E. G., MARZETTA, T. L. "Massive MIMO: Ten Myths and One Critical Question", March 2016.

[58] PAULRAJ, A., PAPADIAS, C. "Space-time processing for wireless communications", *IEEE Signal Processing Magazine*, v. 14, n. 6, pp. 49–83, November 1997.

[59] SARKAR, T. K., JI, Z., KIM, K., et al. "A survey of various propagation models for mobile communication", *IEEE Antennas and Propagation Magazine*, v. 45, n. 3, pp. 51–82, June 2003.

[60] GLANCE, B., GREENSTEIN, L. "Frequency-selective fading effects in digital mobile radio with diversity combining", *IEEE Transactions on Communications*, v. 31, n. 9, pp. 1085–1094, September 1983.

[61] CHEN-NEE CHUAH, TSE, D., KAHN, J., et al. "Capacity scaling in MIMO wireless systems under correlated fading", *IEEE Transactions on Information Theory*, v. 48, n. 3, pp. 637–650, March 2002.

[62] PEEBLES, P. Z. *Probability, random variables, and random signal principles*. McGraw Hill, 2001.

[63] JOSE, J., ASHIKHMIN, A., MARZETTA, T. L., et al. "Pilot contamination and precoding in multi-cell TDD systems", *IEEE Transactions on Wireless Communications*, v. 10, n. 8, pp. 2640–2651, August 2011.

[64] GOPALAKRISHNAN, B., JINDAL, N. "An analysis of pilot contamination on multi-user MIMO cellular systems with many antennas". In: *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications*, pp. 381–385, San Francisco, CA, USA, June 2011. IEEE.

[65] ASHIKHMIN, A., MARZETTA, T. "Pilot contamination precoding in multi-cell large scale antenna systems". In: *2012 IEEE International Symposium on Information Theory Proceedings*, pp. 1137–1141, Cambridge, MA, USA, July 2012. IEEE.

[66] MULLER, R. R., COTTATELLUCCI, L., VEHKAPERA, M. "Blind Pilot Decontamination", *IEEE Journal of Selected Topics in Signal Processing*, v. 8, n. 5, pp. 773–786, October 2014.

[67] NGO, H. Q., LARSSON, E. G. "EVD-based channel estimation in multi-cell multiuser MIMO systems with very large antenna arrays". In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3249–3252, Kyoto, Japan, March 2012. IEEE.

[68] FERNANDES, F., ASHIKHMIN, A., MARZETTA, T. L. "Inter-Cell Interference in Noncooperative TDD Large Scale Antenna Systems", *IEEE Journal on Selected Areas in Communications*, v. 31, n. 2, pp. 192–201, February 2013.

[69] KHANSEFID, A., MINN, H. "On Channel Estimation for Massive MIMO With Pilot Contamination", *IEEE Communications Letters*, v. 19, n. 9, pp. 1660–1663, September 2015.

[70] YANG, S., HANZO, L. "Fifty Years of MIMO Detection: The Road to Large-Scale MIMOs", *IEEE Communications Surveys & Tutorials*, v. 17, n. 4, pp. 1941–1988, 2015.

[71] WIESEL, A., ELDAR, Y. C., SHAMAI, S. "Zero-Forcing Precoding and Generalized Inverses", *IEEE Transactions on Signal Processing*, v. 56, n. 9, pp. 4409–4418, September 2008.

[72] SHI, S., SCHUBERT, M., BOCHE, H. "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization", *IEEE Transactions on Signal Processing*, v. 55, n. 11, pp. 5436–5446, November 2007.

[73] WIESEL, A., ELDAR, Y., SHAMAI, S. "Linear precoding via conic optimization for fixed MIMO receivers", *IEEE Transactions on Signal Processing*, v. 54, n. 1, pp. 161–176, January 2006.

[74] STOJNIC, M., VIKALO, H., HASSIBI, B. "Rate maximization in multi-antenna broadcast channels with linear preprocessing", *IEEE Transactions on Wireless Communications*, v. 5, n. 9, pp. 2338–2342, September 2006.

[75] JOHAM, M., UTSCHICK, W., NOSSEK, J. "Linear transmit processing in MIMO communications systems", *IEEE Transactions on Signal Processing*, v. 53, n. 8, pp. 2700–2712, August 2005.

[76] LO, T. K. Y. "Maximum ratio transmission", *IEEE Transactions on Communications*, v. 47, n. 10, pp. 1458–1461, October 1999.

[77] CHEN, Y., TELLAMBURA, C. "Performance Analysis of Maximum Ratio Transmission with Imperfect Channel Estimation", *IEEE Communications Letters*, v. 9, n. 4, April 2005.

[78] CAIRE, G., SHAMAI, S. "On the achievable throughput of a multiantenna Gaussian broadcast channel", *IEEE Transactions on Information Theory*, v. 49, n. 7, pp. 1691–1706, July 2003.

[79] NGO, H. Q., LARSSON, E. G. "No Downlink Pilots Are Needed in TDD Massive MIMO", *IEEE Transactions on Wireless Communications*, v. 16, n. 5, pp. 2921–2935, May 2017.

[80] NGO, H. Q., LARSSON, E. G., MARZETTA, T. L. "Aspects of favorable propagation in massive MIMO", pp. 76–80, 2014.

[81] GUNNARSSON, S., FLORDELIS, J., VAN DER PERRE, L., et al. "Channel Hardening in Massive MIMO-A Measurement Based Analysis", 2018.

[82] MOLISCH, A., WIN, M. "MIMO systems with antenna selection", *IEEE Microwave Magazine*, v. 5, n. 1, pp. 46–56, March 2004.

[83] SANAYEI, S., NOSRATINIA, A. "Antenna selection in MIMO systems", *IEEE Communications Magazine*, v. 42, n. 10, pp. 68–73, October 2004.

[84] SWINDLEHURST, A. L., JEDDA, H., FIJALKOW, I. "Reduced Dimension Minimum Ber Psk Precoding For Constrained Transmit Signals In Massive Mimo", *IEEE ICASP 2018*, April 2018.

[85] HEATH, R. W., SANDHU, S., PAULRAJ, A. "Antenna selection for spatial multiplexing systems with linear receivers", *IEEE Communications Letters*, v. 5, n. 4, pp. 142–144, April 2001.

[86] GHARAVI-ALKHANSARI, M., GERSHMAN, A. B. "Fast antenna subset selection in MIMO systems", *IEEE Transactions on Signal Processing*, v. 52, n. 2, pp. 339–347, February 2004.

[87] HANIF, M., YANG, H. C., BOUDREAU, G., et al. "Low complexity antenna subset selection for massive MIMO systems with multi-cell cooperation", *2015 IEEE Globecom Workshops, GC Wkshps 2015 - Proceedings*, December 2015.

[88] JOSHI, S., BOYD, S. "Sensor selection via convex optimization", *IEEE Transactions on Signal Processing*, v. 57, n. 2, pp. 451–462, February 2009.

[89] SHAMAIAH, M., BANERJEE, S., VIKALO, H. "Greedy sensor selection: Leveraging submodularity", *Decision and Control (CDC), 2010 49th IEEE Conference on*, pp. 2572–2577, December 2010.

[90] CHEPURI, S. P., LEUS, G. "Sensor selection for estimation, filtering, and detection", *Signal Processing and Communications (SPCOM), 2014 International Conference on*, pp. 1–5, July 2014.

[91] JAMALI-RAD, H., SIMONETTO, A., LEUS, G. "Sparsity-aware sensor selection: Centralized and distributed algorithms", *IEEE Signal Processing Letters*, v. 21, n. 2, pp. 217–220, February 2014.

[92] CHENG, C. *Optimality of Certain Asymmetrical Experimental Designs*, v. 6. Institute of Mathematical Statistics, 1978.

[93] DUA, A., MEDEPALLI, K., PAULRAJ, A. J. "Receive antenna selection in MIMO systems using convex optimization", *IEEE Transactions on Wireless Communications*, v. 5, n. 9, pp. 2353–2357, September 2006.

[94] MAHBOOB, S., RUBY, R., LEUNG, V. C. "Transmit antenna selection for downlink transmission in a massively distributed antenna system using convex optimization", *Broadband, Wireless Computing, Communication and Applications (BWCCA), 2012 Seventh International Conference on*, pp. 228–233, November 2012.

[95] TIBSHIRANI, R. "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[96] FRIEDMAN, J., HASTIE, T., TIBSHIRANI, R. *The elements of statistical learning*, v. 1. Springer series in statistics New York, 2001.

[97] HANIF, M., YANG, H.-C., BOUDREAU, G., et al. "Antenna subset selection for massive MIMO systems: A trace-based sequential approach for sum rate maximization", *Journal of Communications and Networks*, v. 20, n. 2, pp. 144–155, April 2018.

[98] HARVILLE, D. A. *Matrix algebra from a statistician's perspective*, v. 1. Springer, 1997.

[99] MALLAT, S., ZHANG, Z. "Matching pursuits with time-frequency dictionaries", *IEEE Transactions on Signal Processing*, v. 41, n. 12, pp. 3397–3415, December 1993.

[100] GOLUB, G. H., LOAN, C. F. V. *Matrix Computations*. The John Hopkins University Press, Baltimore and London, 1996.

[101] STEWART, G. W. *Matrix algorithms volume 2: eigensystems*, v. 2. Siam, 2001.

[102] GRANT, M., BOYD, S., YE, Y. "CVX, Version 1.21 MATLAB software for disciplined convex programming". 2007.

[103] CHOI, J., SUNG, J., EVANS, B. L., et al. "Antenna Selection for Large-Scale MIMO Systems with Low-Resolution ADCs", *arXiv preprint arXiv:1801.09846*, 2018.

[104] SUNG, J., CHOI, J., EVANS, B. L. "Narrowband Channel Estimation for Hybrid Beamforming Millimeter Wave Communication Systems with One-bit Quantization", *arXiv preprint arXiv:1710.10673*, 2017.

[105] MO, J., SCHNITER, P., HEATH, R. W. "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs", *IEEE Transactions on Signal Processing*, v. 66, n. 5, pp. 1141–1154, 2018.

[106] GOPALA, K., SLOCK, D. T. "Optimal algorithms and CRB for reciprocity calibration in massive MIMO". In: *ICASSP 2018, IEEE International Conference on Acoustics, Speech and Signal Processing, 15-20 April 2018*, Calgary, CANADA, 04 2018.