Democratic and Popular Republic of Algeria
Ministry of Higher Education and Scientific Research

UNIVERSITYOF SAIDA Dr. TAHAR MOULAY SAIDA

Faculty of Technology
Department of Electronic

Dissertation Submitted to the Department of electronic as a partial of Requirements for the Master Degree in Telecommunication

Option: Networks and Telecommunications

Algorithms of detection in Massive MIMO for new wireless communication systems reduction matrix inverse

Presented by:
BOUREGAG Khaoula
SI YOUCEF Imene

Supervised by:
Dr. TAMI Abdelkader

Broad of Examiners:
Dr. BOUYEDO Benamer — *president*
Dr. OUARDI Aissa — *examiner*
Dr. TAMI Abdelkader — *supervisor*

Academic year: 2019-2020

Acknowledgments

First of all, all thanks and praise to our God "God" for granting us the courage and patience to complete this work.

This work was carried out under the supervision of Mr. Tami Abdelkader, so we would like to express our gratitude for making use of our scientific skills.

We warmly thank our families for their moral and material assistance throughout the preparation period

We would also like to thank all the teachers who contributed to our training in the technology department.

Finally, we send our sincere thanks to all the people who helped us, near or far, to develop our humble memory

اهداء

إن الحمد و الشكر لله الذي أعانني على إنجاز هذا العمل و يسر لي أموري وأمدني بالصبر و العزيمة.

اهدي ثمرة عملي إلى الراحل من حياتي الحاضر في قلبي إلى من حصد الأشواك ليحميني ويمهد لي طريق العلم, كان لي السراج الذي أنار درب حياتي وألهمني الصمود والتحدي لبلوغ مقاصدي والدي الحبيب رحمك الله وجعل الجنة مثواك

إلي من كان دعائها سر نجاحي إلى رمز الحب الى سندي في الحياة أمي الغالية حفظها الله وأدامها لي

إلى من هم اقرب أليّ من روحي وبهم استمد عزتي وإصراري أخواتي حفظهم الله: كريمة, سماح , هند , ريم , أمال , منال

إلى كتاكيت عائلتي: ريهام و قاسم

إلى أختي التي لم تلدها أمي وزميلتي في المذكرة " سي يوسف إيمان"

إلى كل الذين أحبهم

خولة

اهداء

أيام مضت من عمرنا بدأناها بخطوة و ها نحن اليوم نقطف ثمار مسيرة أعوام كان هدفنا فيها واضحاً و كنا نسعى في كل يوم لتحقيقه و الوصول اليه مهما كان صعباً و ها نحن وصلنا و بيدينا شعلة علم و سنحرص عليها كل الحرص حتى لا تنطفئ

اولا الحمد لله الذي منّ عليّ بأن أكون في مكاني هذا أكتب هاته الكلمات و منحني فرصة شكر أشخاص غالين على قلبي

أهدي ثمرة هذا العمل المتواضع أولا الى القلب الحنون من كانت بجانبي بكل المراحل التي مضت من تلذذ و معاناة و كانت شمعة تحترق لتنير دربي امي الغالية

و كما ان الفضل كل الفضل لوالدي رحمة الله عليه الذي علمني الصعود و عيناه تراقبني و ساعدني بكل مراحل حياتي, نم قرير العين فإبنتك قد تخرجت وحققت حلمك سأسعى جاهدة لأكون فخرك و عزك في الدنيا و الآخرة

"الى امي الثانية منبع العطف والحنان " زوجة خالي

الى من شاركوني حنان و عطف الوالدين اخواتي حفظهم الله سمية،رحمة،سولاف،حليمة،حفيظة،كلتومة،خولة،سعاد،جميلة

الى براعم حياتي قطرالندى،ابراهيم،امينة

و اخص بالشكر الإنسانة التي لن توفيها حقها الكلمات اختي وزميلتي في هذه المذكرة (بورقاق خولة) على تعاونها معي ولا أنسى زملاء الدراسة كل واحد باسمه

وايضا اهدي تخرجي لكل من كان له بصمة في حياتي فلكم كل الود و الاحترام

ايمان

Abstract:

The need to accommodate more users at higher data rates with better reliability while consuming less power has forced the birth of a new generation of 5G mobile communication based on new proposed techniques. Among these techniques a Massive MIMO (Multi-Input, Multi-Output) system which is relied on multipath benefit to enhance system capability. The objective of this work is to study some algorithms as Principal Component Method and Krylov Subspace that can reduce the complexity of detectors of zero forcing (ZF) and minimum mean-square error (MMSE) who base on inversing the channel matrix of a massive MIMO in 5G system.

Keywords: Mobile Network; 5G; Massive MIMO; Detectors, Principal Component, Krylov Subspace.

الملخص:

أدت الحاجة إلى استيعاب المزيد من المستخدمين بمعدل تدفق أعلى ودقة أفضل مع استهلاك اقل للطاقة، إلى توليد جيل جديد للاتصالات النقالة،الجيل الخامس الذي يعتمد على تقنيات حديثة، من أهمها تقنية massif MIMO (متعدد المداخل، متعددة المخارج)المكثفة، تعتمد هذه التقنية على الاستفادة من المسارات المتعددة لتحسين سعة النظام. الهدف من هذا المشروع هو دراسة بعض الخوارزميات كطريقة المكون الرئيسي (Principal Component Method)والفضاء الجزئي لكيرلاف (Subspace Krylov)التي يمكن أن تقلل من تعقيد أجهزة الكشف للتأثير الصفري (ZF) والحد الأدنى من الخطأ التربيعي (MMSE) الذين يعتمدون على عكس مصفوفة قناة MIMO الضخمة في نظام الجيل الخامس.

كلمات المفاتيح :شبكة المحمول،الجيل الخامس، MIMO المكثفة، أجهزة الكشف، طريقة المكون الرئيسي، والفضاء الجزئي لكيرلاف

Résumé :

Le besoin d’accueillir plus d’utilisateurs à des débits de données plus élevés avec une meilleure fiabilité tout en consommant moins d’énergie a imposé la naissance d’une nouvelle génération de communication mobile la 5G basant, parmi les nouvelles techniques proposées, la technique massive MIMO (Multi-Input, Multi-Output), ce système est appuyé sur le bénéficier de trajet multiple pour améliorer la capacité de système. L'objectif de ce projet est d'étudier certains algorithmes comme la méthode de composant principal et le sous espace de Krylov qui peuvent réduire la complexité des détecteurs comme le forçage nul (ZF) et l'erreur quadratique moyenne minimale (MMSE) qui reposent sur l'inverse de la matrice de canaux d'un MIMO massif dans un système 5G

Mots clés : Réseau mobile ; 5G ; MIMO massif, Détecteurs, Principal Component, Krylov Sous-Espace

Index

Chapter III: Algorithms of detection in Massive MIMO System

List of figures

Chapter I: New wireless communications systems

Chapter II: MIMO Massive

List of figures

List of tables

Shortcuts

1G: The first Generation

2G: The second Generation

3G: The third Generation

4G: The fourth Generation

5G: The fifth Generation

RAN: Radio Access Network

BS: Base Station

MIMO: Multiple-Input Multiple-Output

NB: Narrow Band

HeNB: Home eNodeB

3GPP: Third Generation Partnership Project

CSG: Closed Subscriber Groups

MF: Matched Filter

BS: Base Station

IoT: Internet of Things

RAT: Radio Access Technology

FD-MIMO: Full-Dimension MIMO

MU-MIMO: Multi User MIMO

3D: Three-dimensional

NFV: Network Function Virtualization

SDN: Software-Defined Networking

NR: New Radio

NOMA: Non-Orthogonal Multiple Access

UE: User Equipment

CP: Cyclic Prefix

Shortcuts

OFDM: Orthogonal Frequency Division Multiplexing

NB-IoT: Narrow Band Internet of Things

LTE: Long Term Evolution

RB: Resource Block

CBW: Channel Band Width

SP: Subcarrier

BG: guard Band

SISO: Single Input Single Output

SIMO: Single entry Multiple exit

MISO: Single output with multiple inputs

MIMO: Multiple input Multiple output

LoS: Line-of-sight

CSI: Channel State Information

SNR: Signal-to-noise ratio

m-MTC: Massive Machine-Type Communications

TDD: Time Division Duplexing

FDD: Frequency Division Duplexing

MMSE: Minimum Mean Square Error

SE: Spectral Efficiency

SU-MIMO: Single-User MIMO

AWGN: Additive White Gaussian Noise

MU-Massive MIMO: MultiUser Massive MIMO

CDMA: Code Division Multiple Access

ZF: Zero Forcing

MRC: Maximum Ratio Combining

ML: Maximum-likelihood

SD: Sphere decoder

Shortcuts

SINR: Signal-to-interference ration

NS: Neumann Series

GS: Gauss–Seidel

SOR: Successive Over Relaxation

JA: Jacobi

MSE: Mean-Square Error

QMR: Quality Management Representative

PC/PCA: Principal Components / Principal ComponentsAnalysis

KS/KST: Krylov Subspace / Krylov Subspace Techniques

GMRES: Generalized Minimal Residual method

General Introduction

Future wireless networks will have to deal with a substantial increase of data transmission due to a number of emerging applications that include machine-to-machine communications and video streaming [1]- [4]. This very large amount of data exchange is expected to continue and rise in the next decade so, presenting a very significant challenge to designers of fifth generation (5G) wireless communications systems [4].Amongst the main problems are how to make the best use of the available spectrum and how to increase the energy efficiency in the transmission and reception of each information unit. 5G communications will have to rely on technologies that can offer a major increase in transmission capacity as measured in bits/Hz/area but do not require increased spectrum bandwidth or energy consumption.

Massive MIMO signal detection is the key technology of next generation wireless communication (such as 5G). In massive MIMO signal detection, there are many algorithms that can implement the signal detection. Generally, these algorithms can be divided into linear detection algorithm and nonlinear detection algorithm according to different calculation methods. Although the linear detection algorithm is less accurate than the nonlinear detection algorithm, it is still an effective signal detection method of massive MIMO system in some cases due to its low complexity. Massive MIMO is destined to provide great and improved user experience, delivery of new revenue generated exciting mobile services. Consequently, massive MIMO would remain a strong competitor in the next decade for both developed and emerging markets. A significant research dedicated to the receiver's design has been proposed. In this dissertation, a review of various detection techniques for massive MIMO systems is provided. Although linear detectors suffer from mediocre performance, the ZF method and the MMSE method are found to play a crucial role in the receiver design due to their relative simplicity. They are also used in the initialization and pre-processing for other detectors. may achieve a promising balance between performance and complexity.

In massive MIMO system the huge number of transmitted and received antennas increase considerably the complexity of detection algorithm like ZF and MMSE detectors. This work comes to present some method of reducing algorithms of detection in Massive MIMO such as Principal Component method (PC) and Krylov Subspace.

This document is structured around four chapters:

First chapter will present a description of the next generation of mobile network 5G.

Second chapter will give the evolution of massive MIMO.

Third chapter will deal with algorithm of detection in massive MIMO system also the effect of the Principal Component method (PC) and Krylov Subspace to reduce the complexities of massif MIMO detectors

The last chapter will be dedicated to show of different simulations of reduced complexity algorithms using different parameters.

Chapter I

I.1 Introduction:

The mobile network is a wireless technology capable of providing a voice and / or data network by radio transmission. The mobile phone is one of the best-known applications in the mobile network. Previously, circuit switching was used to transmit voice over a network, and then we used both circuit switching and packets witching for voice and data. Currently, packets witching are only used. This is how the spectrum went from 1G (the first generation) to 4G (the fourth generation). Today and in the years to come, wireless networks need to be improved to meet the demand for increased data throughput, improved capacity, reduced latency and quality of service. Service, we are in the 4th generation of wireless communication, research is being carried out to develop new standards for the next generation beyond 4G, with the growing demand of subscribers, and 4G will be definitively replaced by 5G at the using advanced technologies. The purpose of this chapter is to describe the context of mobile applications and the main challenges associated with it in order to present the arrival of the fifth generation.

I.2 Evolution of mobile communications:

In just 20 years, mobile phone networks have undergone a profound transformation from the second generation (2G) (voice communication), to the third generation (3G) (voice communication and data transfer), then to the fourth generation (4G) (broadband communication and transfer). Indeed, from a telephone originally designed to carry out a voice conversation between two users without providing any other service than the simple sending / receiving of text messages, the modern Smartphone has today become a real portable data center giving access to a multitude of services and applications (camera, internet browser, games, etc.). This generalization of internet access via the Smartphone and the arrival of video calls, push for an even higher data rate.[1]

Figure (I.1): Evolution of mobile networks

Since 1979 with the deployment of the first generation (1G) of wireless networks, mainly analog, a new technology has been born every 10 years **Figure (I.1)**. The second generation (2G) of mobile networks and the digital switch over with the creation of the GSM standard has introduced new data transfer services such as SMS (Short Message Service). However, 2G could not yet respond to the demand for internet access from mobile phones. This motivated the development of 3G, which arrived on the marketing 2001 with the first Smartphone's. The nin 2009, data transfer rates much higher than those of 3G could be achieved with the definition of 4G allowing video calls and much larger file transfers. However, the 4G network cannot meet the growing demand for the number of connected objects. This is why the mobile phone industry has chosen to develop the fifth generation (5G) network to provide a technical solution to the problem facing 4G today.

The 5G concept combines both an evolution of existing mobile networks to meet future data transfer demands, but also a revolution with the creation of a new, more competitive communication technology, i.e. a network which will be more efficient and cheaper.

I.3 5G communication system architecture:

In 5G communications, the adoption of a dense heterogeneous architecture, comprising macro cells and small cells, is one of the most promising low-cost solutions that will allow5G networks to meet the industry’s capacity growth needs and to provide a uniform connectivity experience on the end user’s side. Based on the latest literature, we consider in **Figure (I.2)** that a potential 5Gcommunications architecture in a macro cell scale, as depicted in, will include the base station (BS), equipped with large antenna arrays, as well as additional large antenna arrays of the BS geographically distributed over the macro cell network. The distributed large antenna arrays will play the role of small-cell access points supporting multiple Radio Access Network (RAN) protocols for a wide range of underlying access network technologies (2G/3G/4G). Moreover, mobile users in the outdoor environment will collaborate with each other to form virtual large antenna arrays. The virtual large antenna arrays, together with the distributed large antenna arrays (i.e. small-cell access points) of the BS, will construct virtual massive MIMO (Multiple-Input Multiple-Output) links in the small cells. The small-cell access points rely on reliable backhaul connectivity over optical fibers. [2]

Furthermore, the buildings located in the 5G macro cell area will also be equipped with large antenna arrays installed outside of the building. Thus, every building will be able to communicate with the BS of the macro cell directly or with the distributed large antenna arrays of the BS. Besides, in every building, the outside installed large antenna arrays will be connected via cable to the wireless access points inside the building communicating with indoor users. [2]

Additionally, the Home eNodeB (HeNB) reference architecture, defined by 3GPP (Third Generation Partnership Project) in references in order to construct femtocells, is very promising for the upcoming5Gcommunications networks. This is because the HeNB femtocell provides an effective solution to address the increasing demand for data rates. In particular, an HeNB femtocell is a low-power and low-range access point mainly used to provide indoor coverage for Closed Subscriber Groups (CSG). HeNB femtocells offload the macro cell network and provide broadband IP backhaul connection to the mobile operator’s network through the subscribers residential Internet access. A number of HeNB femtocells may be grouped and addressed to a gateway, reducing the number of interfaces linked directly with the mobile operator’s core

Network. This gateway is a mobile network operator’s equipment which is usually located physically on mobile operator premises.[2]

Moreover, the concept of mobile femtocell (MF_matchedfilter_emtocell) combines the mobile relay concept with femtocell technology to accommodate high-mobility users, such as users on public transport (e.g. trains and buses), and even users in private cars. MF emto cells will be small cells installed inside vehicles to communicate with users within the vehicles.

Also, large antenna arrays will be installed outside the vehicles to enable communication with the BS of the macro cell directly or with the distributed large antenna arrays of the BS. [2]

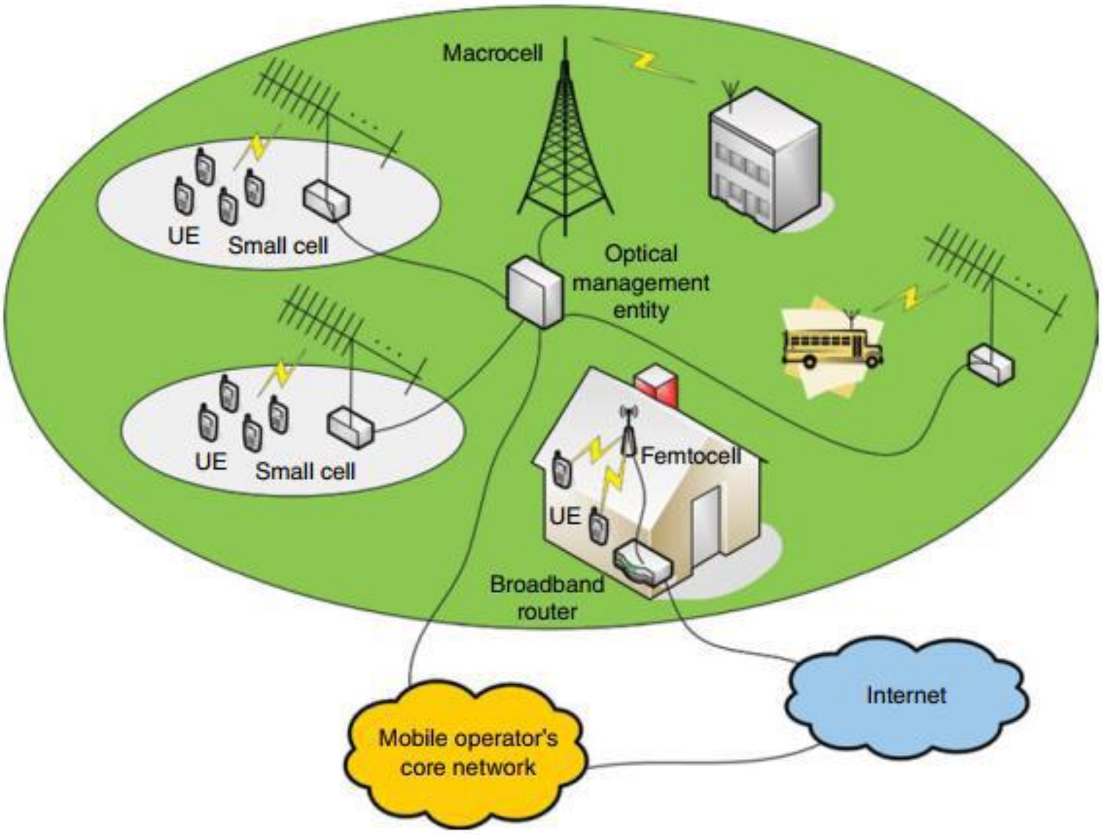


Figure (I.2): 5G system architecture

I.4 Cellular networks:

A cellular network is a telecommunications system which must meet the constraints of the subscriber's mobility in the network, by the extent of the network and by the radio waves allocated to it.

A cellular network system covers the set of infrastructures specially intended for communication routing equipment to mobiles or radio waves, in the case of a cellular network serving as a link between the subscriber's terminal and the operator infrastructure. [3]

I.4.1 Cellular radiotelephony:

A formerly analog and now digital mobile radio system providing all the services offered by the fixed network, plus that of mobility: possibility of maintaining a communication while on the move (handover) and the possibility of calling and being called when you are abroad (international roaming) ....

I.4.2 Cell concept:

Cellular telephony brings together all two-channel radios, one for transmission and one for reception, avoiding likely interference.

The cellular concept makes it possible to reach unlimited capacity by means of a large number of radio stations, each of which covers a geographical area called "cell".

This concept consists of dividing a territory into cells, each of which is covered by a radio station or base station (BTS) of the network. And thus, the reuse of the same station in order to avoid interference phenomena on the useful signal received by the mobile terminal for the base station.[4]

I.4.3 Definition of the Cell:

We call cell, a geographical area covered by antennas on which there is the availability of a given transmission channel (beacon channel), consisting of an electric radio channel characterized by a given frequency or a couple of given frequencies depending on the insured services.

The cells are arranged adjacent to each other and can cover a radius varying from 5 to 20 km, that is to say that they can serve subscribers located in a circle of 10 to 50 km in diameter. The cell plays the role of interface between the mobile and the cellular exchange. It therefore performs the following functions: [4]

- Assignment of communication channels to mobiles.
- Permanent emission of signaling.
- Communication supervision

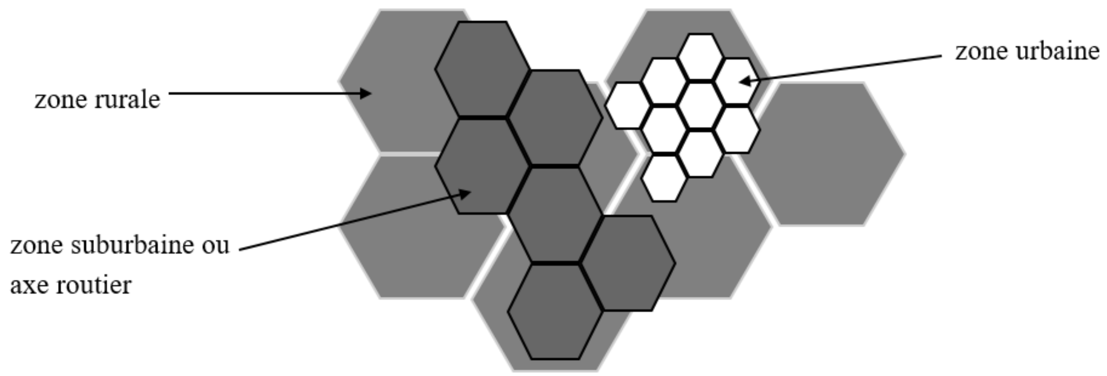


Figure (I.3) : Cell size

I.5 5G goals:

The main manufacturers and operators of the wireless communications sector are currently developing the objectives and standards for the fifth generation of mobile networks (5G). The standardization stage started with indifferent consortia of operators and manufacturers (3GPP, NGMN, etc.) will make it possible to set up regulations by 2020. The first objective to be achieved concerns the up and down speed allocated to each user. To ensure high speed internet access from a smart phone or tablet. In a context of use with little or no mobility, the target speed at the edge of the cell per user is greater than 100 Mbit / s and the maximum speed per user must exceed 10 Gbit / s. In the case of use with high mobility (for example, in the case of communications between vehicles), the improvement of the architecture of the network should in particular make it possible to reduce the latency of the communication.

The objective is to obtain a transmission delay of less than 10 ms. For other IoT (Internet of Things) applications (telemedicine, security, etc.), a high level of reliability will also be required. Major changes to the network architecture and the introduction of new wireless technologies will be required in today's 2G / 3G / 4G networks to achieve these goals. The cost and the energy consumption of the components of this new network and the associated mobile terminals will be decisive points in order to reach an economically viable solution. [4]

I.6 Targeted applications for the 5G network:

5G will provide wireless connectivity for a wide range of new applications shown in **Figure (I.4)**: 5G will penetrate each of the elements of our future society and create a multidimensional user-centered information ecosystem. It will break the limitation of time and space to allow an immersive and interactive user experience. 5G will also shorten the distance between humans and things, and implement seamless integration to achieve asy and intelligent interconnection between people and all things. 5G allows us to realize the vision: "Information is at hand, and everything will stay in touch."

A large number of use cases have been proposed by different organizations. The mobile internet and the Internet of Things (IoT) are the two main drivers of the future development market for mobile communications, and will trigger a wide range of use cases. [5]

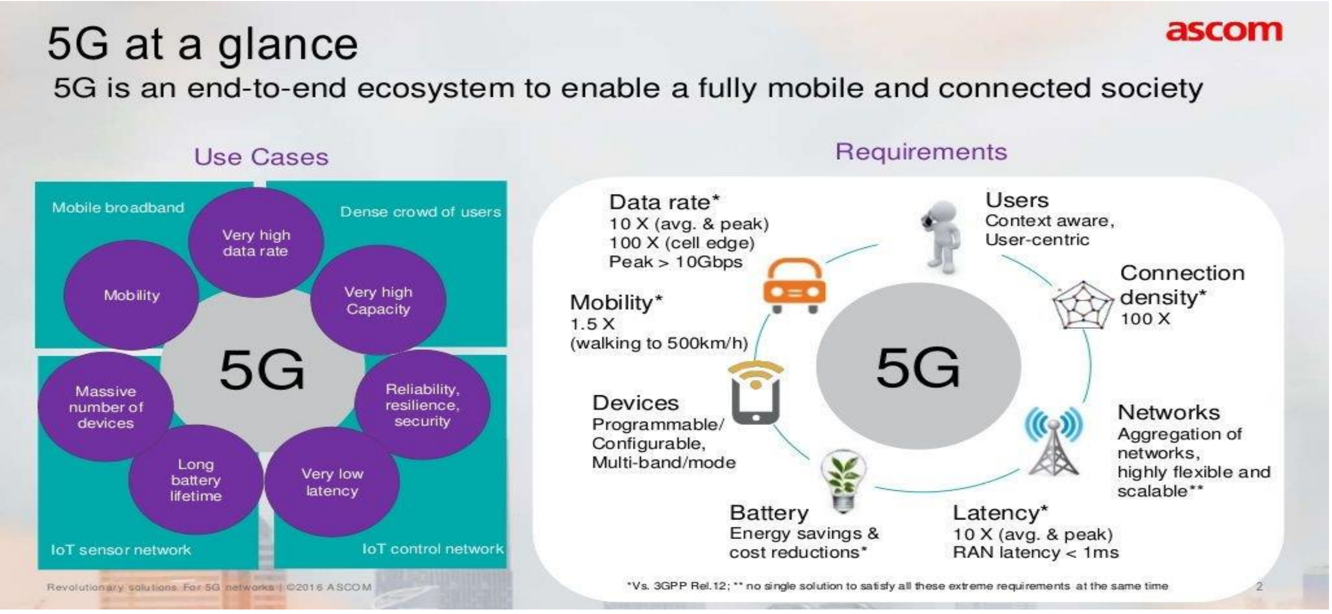


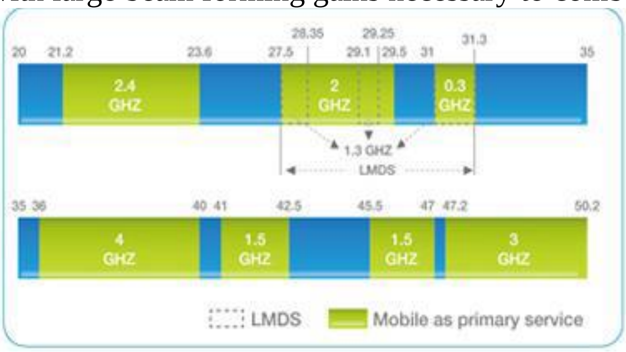
Figure (I.4): Application of 5G

I.7 5G Enabling technologies:

To provide specifics to support the application requirements, 5G enabling technologies are necessary for an ultra-broadband network.

I.7.1 Millimeter wave system:

The mmWave band from 20~50 GHz alone includes 10 times more available bandwidth than the entire 4G cellular band, as illustrated **Fig (I.5.a/b)** Therefore, the mmWave band can support higher data rates required in future mobile broadband access networks.[6]
The small wavelength in mmWave frequency allows design and deployment of massive antenna arrays with large beam forming gains necessary to combat the large propagation loss in the mmWave band.[6]



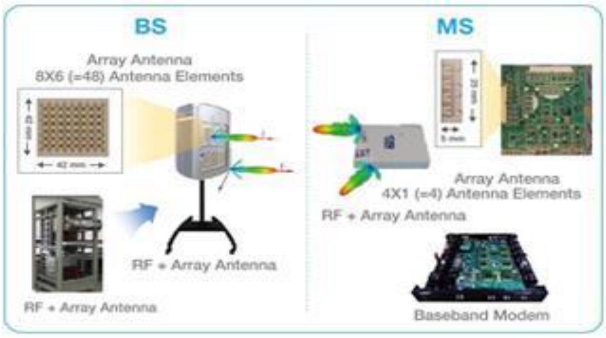


Figure (I.5.a): Potential Bands in 20-50 GHz (US) **Fig (I.5.b):** Adaptive Pencil Beam forming

I.7.2 Mutli-RAT (Multi- Radio Access Technology):

Utilization of large system bandwidth is considered as an effective method to significantly increase per-user throughput and overall system capacity. Finding frequency bands with sizable available bandwidth is therefore one of the key challenges for the 5G system. In this context, the importance of enabling the mmWave band to be utilized in the 5G system is continuously growing. [6]

In parallel with the research on the mmWave, industries are also putting great effort into the evolution of LTE (Long Term Evolution) and the development of a new Radio Access Technology (RAT) on the below 6 GHz. Compared to the RAT (Radio Access Technology) on the mmWave band, the RAT (including LTE) on the below 6 GHz can support large cell coverage and stable wireless links so that it is suitable to send control signals, for instance, messages for paging and handover. Therefore, we expect that interworking and integrating multiple RATs on the different frequency bands into the 5Gsystem is beneficial with respect to both capacity and robustness. We illustrate in **Figure (I.6)** such a system where the RAT on the below 6 GHz is used for exchanging the control information to maintain connections between eNB and users while the mmWave cells support gigabit data rate services.[6]

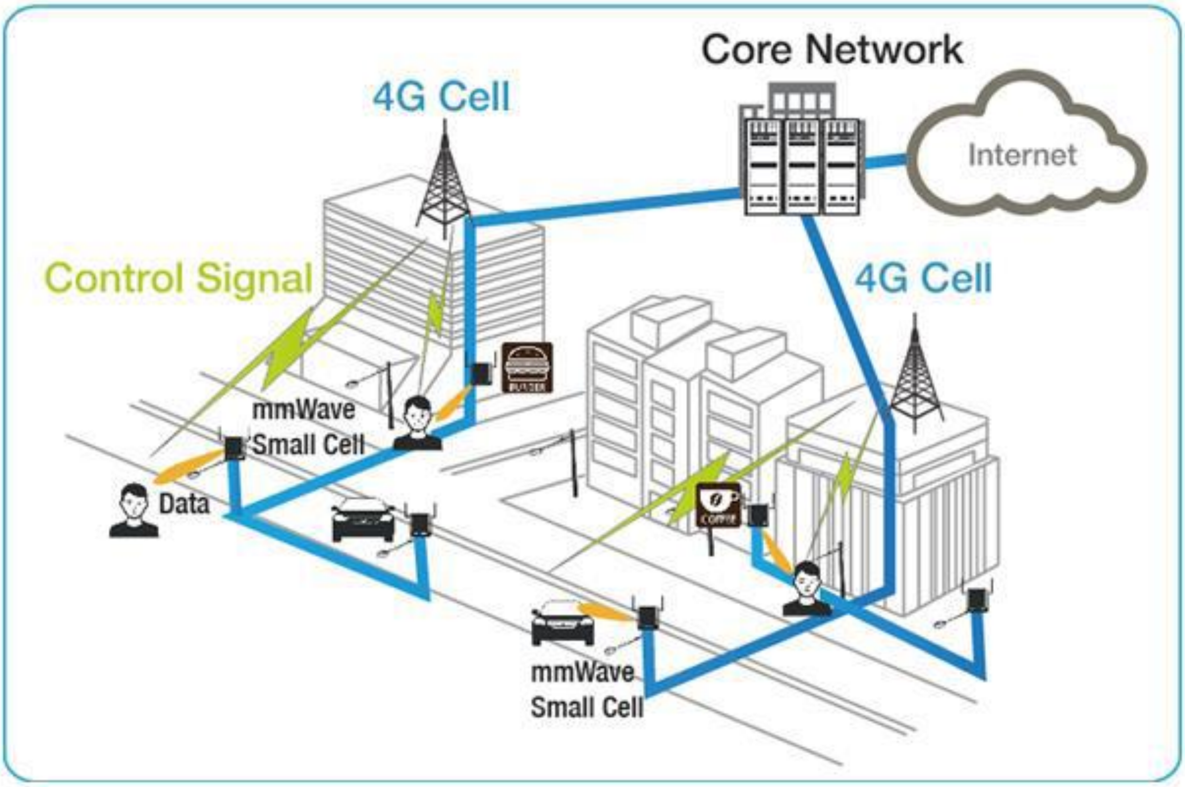


Figure (I.6): Overlaid Network of mmWave Small Cell integrated with the Underlay 4GSystems

I.7.3 Advanced MIMO:

One promising technology for meeting the future demands is massive multiple antennas with multiple-input and multiple-output (MIMO) transmission/reception.

In practice, depending on the operating frequency and form factor requirements of a BS, there is a limit on the number of antennas that can be supported at the BS. For the typical 4Gsystem with frequency bands of 2.5 GHz, fitting 32 antenna elements require up to 1.9 m width, which is not practical for many BSs that have only limited room on the tower. This practical limitation in one-dimension array antenna has motivated Full-Dimension MIMO (FD-MIMO) cellular communication systems, which can place a large number of active antenna elements in a two-dimensional grid at the BSs. [6]

FD-MIMO system can support high-order MU-MIMO through 3D (three-dimension) beam forming algorithms that fully exploit the elevation and azimuth dimensions, thereby generate improved systemthroughput. [6]

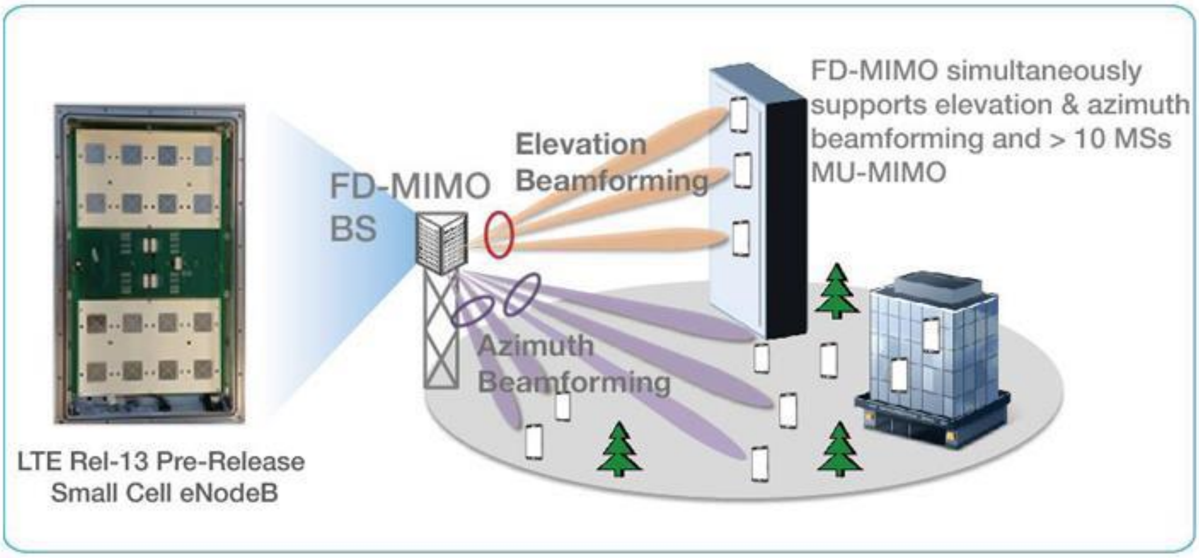


Figure (I.7): Example of FD-MIMO Deployment

I.7.4 Advanced Network:

In order to fulfill key requirements of 5G (such as latency and the large number of simultaneous connections), technologies at the radio access level should be complemented by developments at the system architecture level from the network point of view. A new 5Gnetwork will have to evolve towards a distributed and flat architecture, in order to support the increased data rate facilitated by new 5G radio access technologies. [6]

In addition, 5G network would enable operators to build diverse set of business models and services; flexibility is another key requirement of 5G network architecture. Software-Defined Networking (SDN) and Network Function Virtualization (NFV) provide promising examples of programmable design technologies for realizing a flexible 5G network architecture which enable operators' multi-service adaption of network functions to support a variety ofservices. [6]

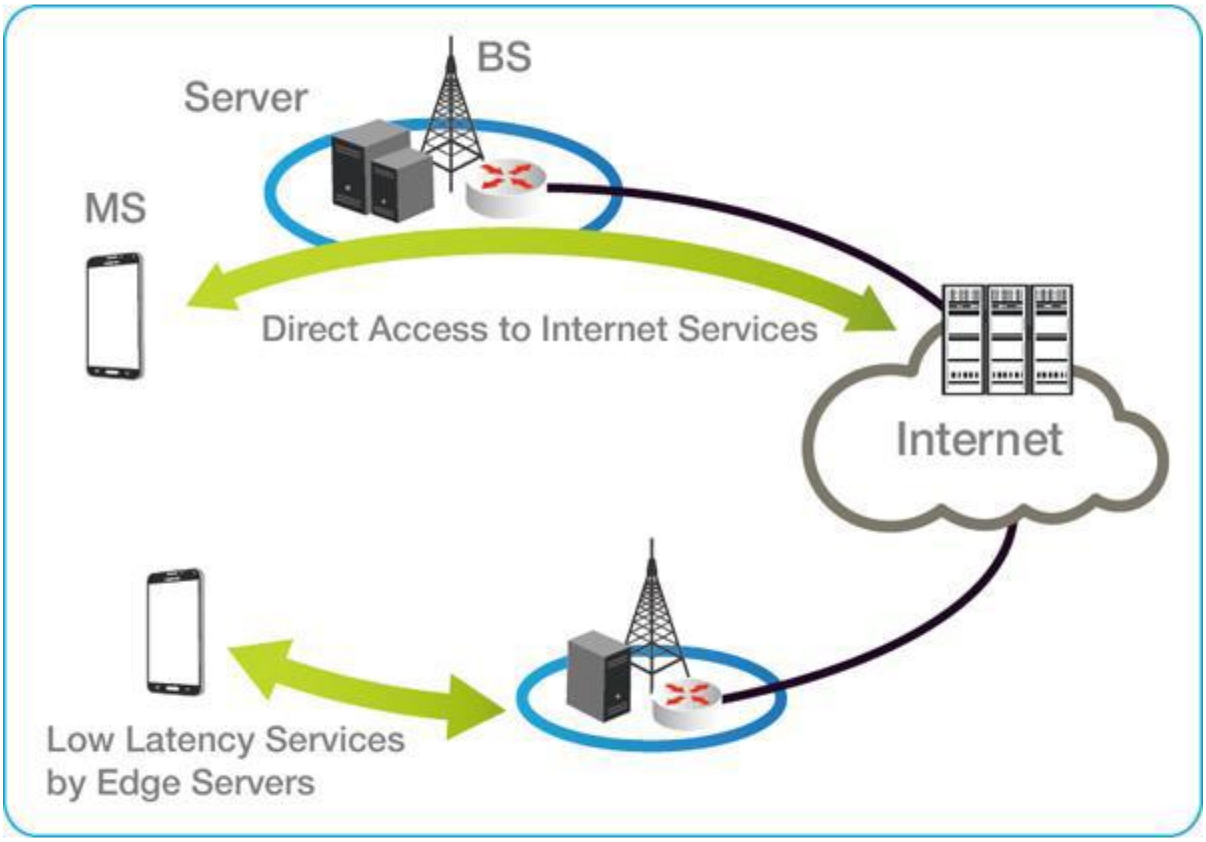


Figure (I.8): 5G Flat Network Architecture

I.8 5G: New Radio (NR):

Meeting the requirements mentioned requires radical changes in the network model in addition to disruptive innovations. In this context, 5G networks can use a wide range of new technologies. This allows a jump in performance that overshadows its predecessors. These innovations will affect the transmission and design of the physical layer in addition to introducing upheavals in the upper layers of the network. In fact, 5G New Radio (NR) will use many key technologies to reach new levels of performance and efficiency. Combinations of these will expand the importance of mobile communications and allow them to play a central role in a world of changing use cases. Among the potential innovations in the 5G physical layer, we can cite: [7]

• Communications in the millimeterwave range.

• Massive MIMO communications.

• Non-Orthogonal Multiple Access (NOMA).

• Full-duplex wireless communications.

• Carrier aggregation and Multi carrier modulations.

• Greater spectrum.

• Sidelink communication.

• New heterogeneous waveform and numerology (OFDM_Orthogonal Frequency Division Multiplexing).

I.8.1 NR deployment scenarios:

To ensure backward compatibility with LTE / LTE-A, the NR architecture is required for close interworking with LTE / LTE-A. For this requirement, the LTE / LTE-A and NR cells may have different coverage (**Figure (1.9)**) or the same coverage, and the following deployment scenarios are feasible: [8]

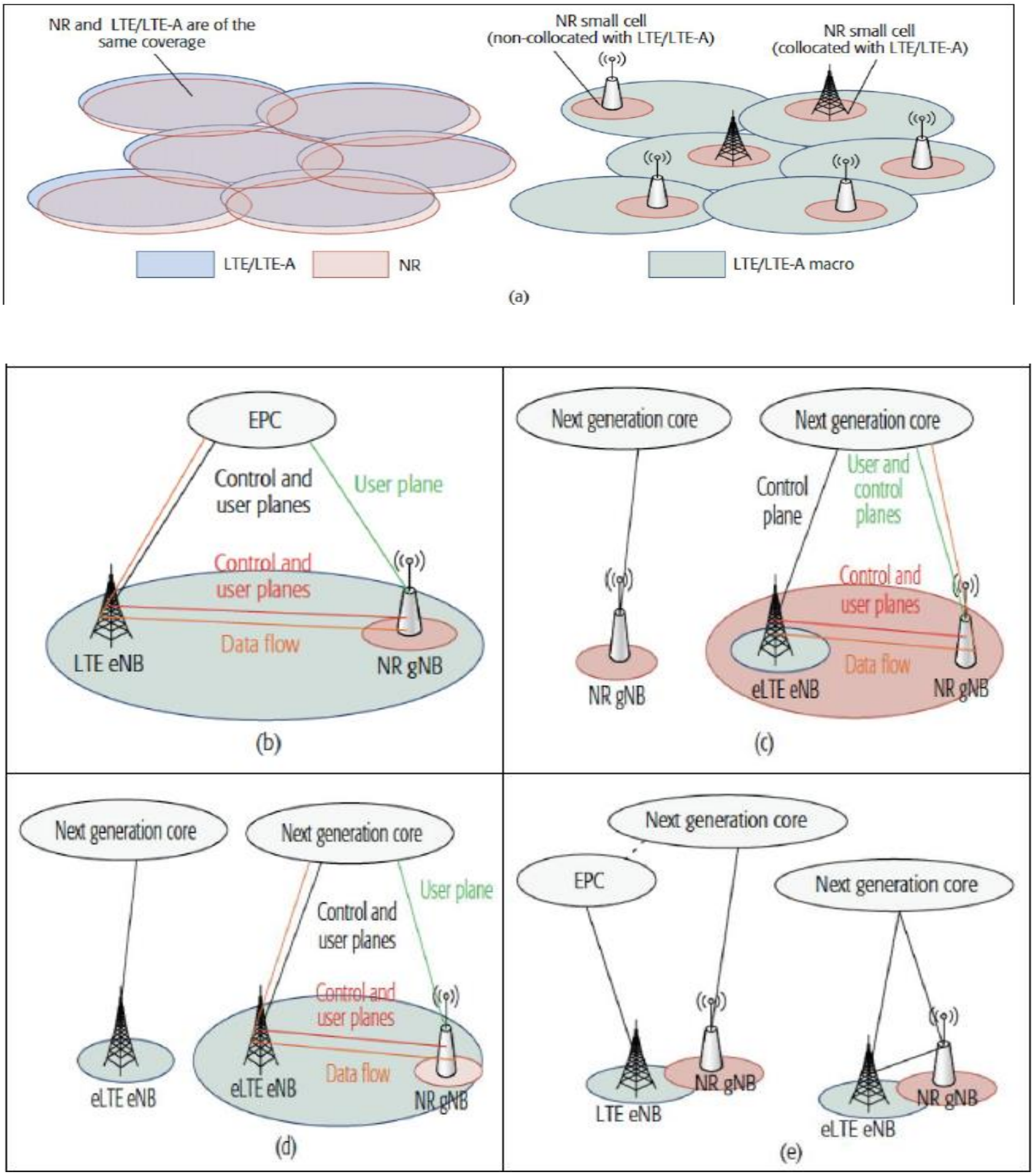


Figure (1.9): NR deployment scenarios. [8]

A LTE/LTE-A eNB connects to the EPC, and a NR gNB connects to the next generation core, to support handover between eNB and gNB. An eLTEeNB can also connect to the next generation core, and handover between eNB and gNB can be fully managed through the next generation core.[8]

I.8.2 Numerologies:

The above scenarios reveal a heterogeneous deployment of NR with different coverage. Taking into account the mobility of user equipment (UE) up to 500 km / h, several lengths of cyclic prefix (CP) must be adopted in NR. In practice, the carrier frequency and the bandwidth of the subcarrier can also affect the length of the adopted CP. Therefore, there can be multiple combinations of physical transmission parameters in NR, such as subcarrier spacing, duration of orthogonal frequency division multiplexing (OFDM) symbols, lengths CP, etc. These physical transmission parameters are collectively referred to as numerology in NR.

In NR, transmitters and receivers can benefit from a wider bandwidth in the high frequency bands. In this case, the subcarrier spacing can be extended (greater than 15 kHz, as adopted by LTE / LTE-A, and possibly up to 960 kHz). In addition, the high carrier frequencies are also vulnerable to the Doppler effect and a large spacing of the subcarriers.

Can help reduce interference between carriers. On the other hand, NR should also support a low spacing of the subcarriers, such as 3.75 kHz, as allows the narrowband internet of things (NB-IoT), to benefit from a better performance. Energy in the low frequency bands. Consequently, the subcarrier spacing's in NR are scalable as a 15 kHz subset or superset. The possible spacing's of the subcarriers can be 15 kHz x 2μ, where μ can be a positive / negative integer or zero. For each subcarrier spacing value, multiple CP lengths can be inserted to accommodate different levels of inter-symbol interference at different carrier frequencies and mobility. [5] The NR Numerology types are summarized in the following table. [12]

μ	$\Delta f = 2\mu.15$ [kHz]	Préfixe cyclique (CP)
0	15	Normal
1	30	Normal
2	60	Normal, Élargie
3	120	Normal
4	240	Normal

Table (I.1) : Supported transmission numerologies [9].

I.8.3 Frame structure :

The downlink and uplink transmissions are organized in 10 ms duration frames, **Figure (I.10)**, each divided into ten 1 ms subframes.

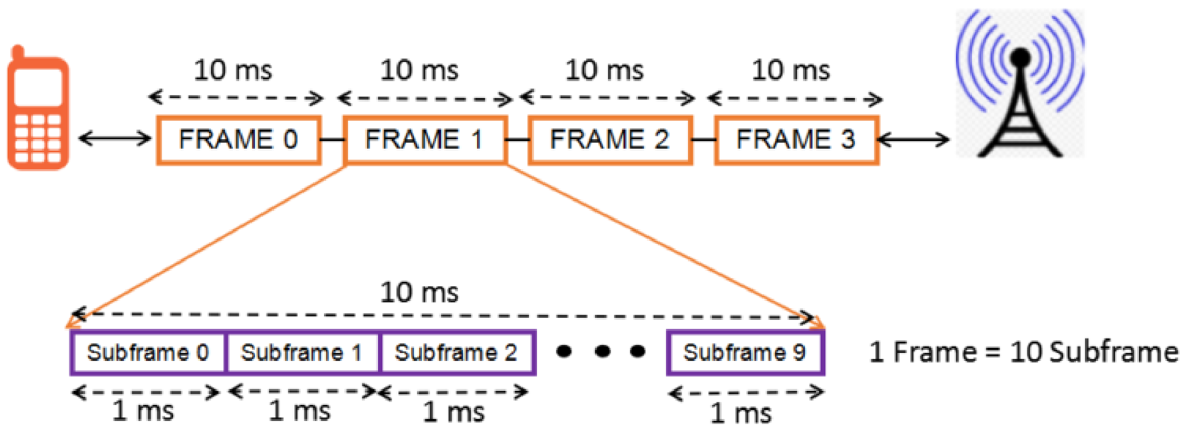


Figure (I.10): Organization of frames in 5G-NR [10]

The slot length varies depending on the spacing of the subcarriers. The general trend is that the slot length decreases as the spacing of the subcarriers widens, **Figure (I.11)**. This trend stems from the nature of the OFDM.

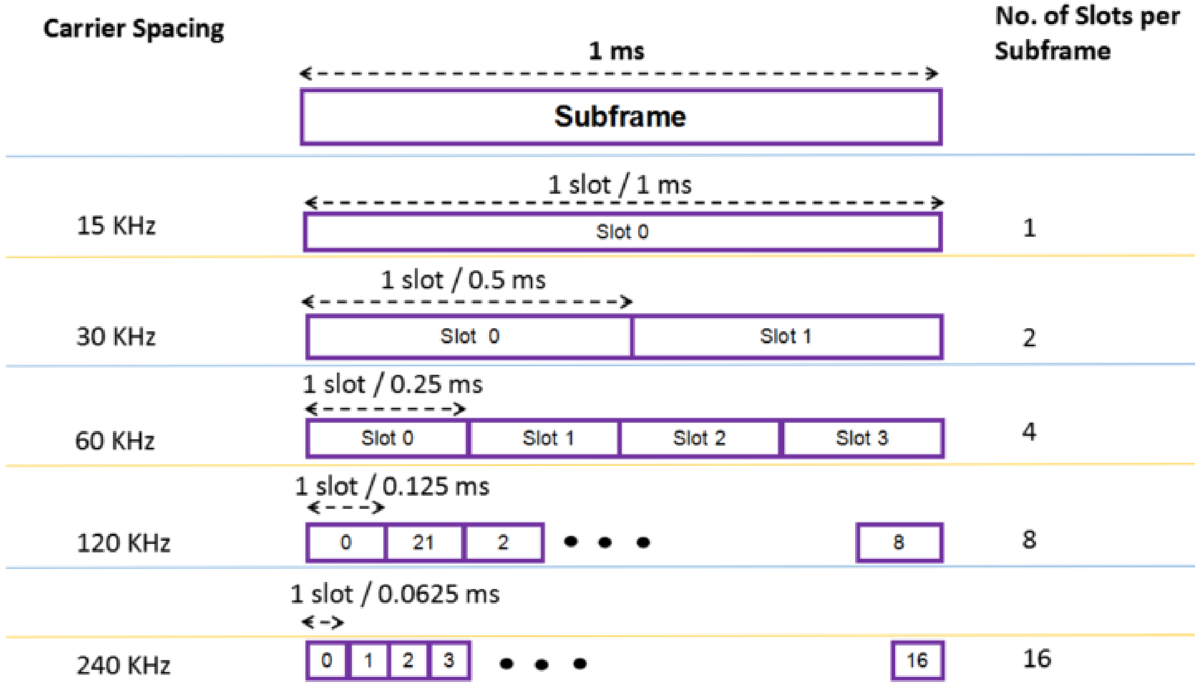


Figure (I.11): Length of slots for each numerology [10]

I.8.4 NR Resource Block:

A NR resource block (RB) contains 14 symbols in the time domain and 12 subcarriers in the frequency domain. In LTE band, the bandwidth of RB is fixed at 180 KHz but in NR, it is not fixed and depends on the spacing of the subcarriers. **Figure (I.12)** shows the SP spacing's for each numerology. [11]

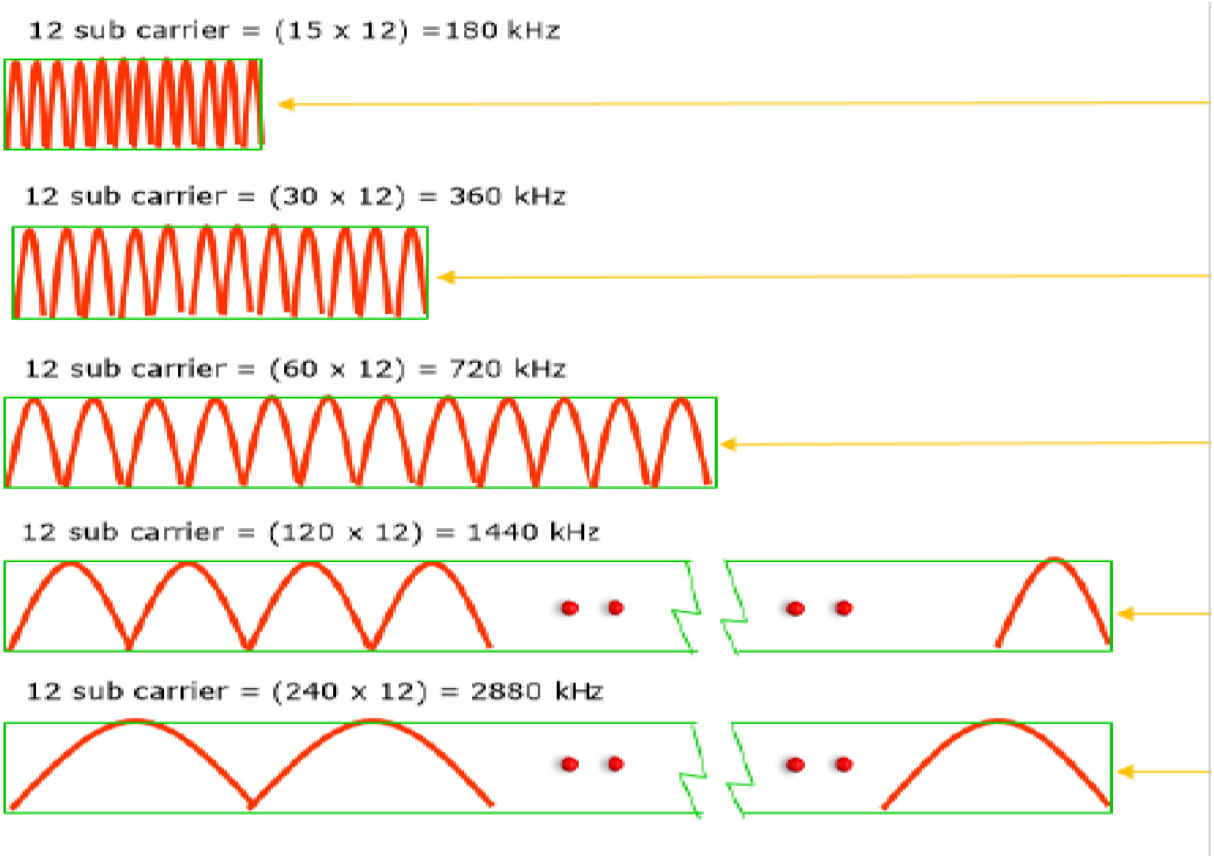


Figure (I.12): Numerology – Subcarrier spacing [11]

I.8.5 NR channel band width:

The NR should operate with a channel width of 100 MHz for the lower bands <6 GHz and a band of 400 MHz in the millimeterwave range. NR is designed to provide higher bandwidth efficiency, up to 99%, or around 90% in LTE (10% wasted in guard band).

Each numerology defined the number of RB, having knowledge of the width of a block of resources, it is possible to calculate the bandwidth of the channel such as: [9]

$$C_{BW} = N_{RB} \times N_{SP} \times \Delta f + 2\text{xBG}$$

Where:

-C_{BW}: channel band width

-N_{RB}: Number of block resources

- N_{SP}: Number of subcarriers = 12

-**BG** : the Guard Band

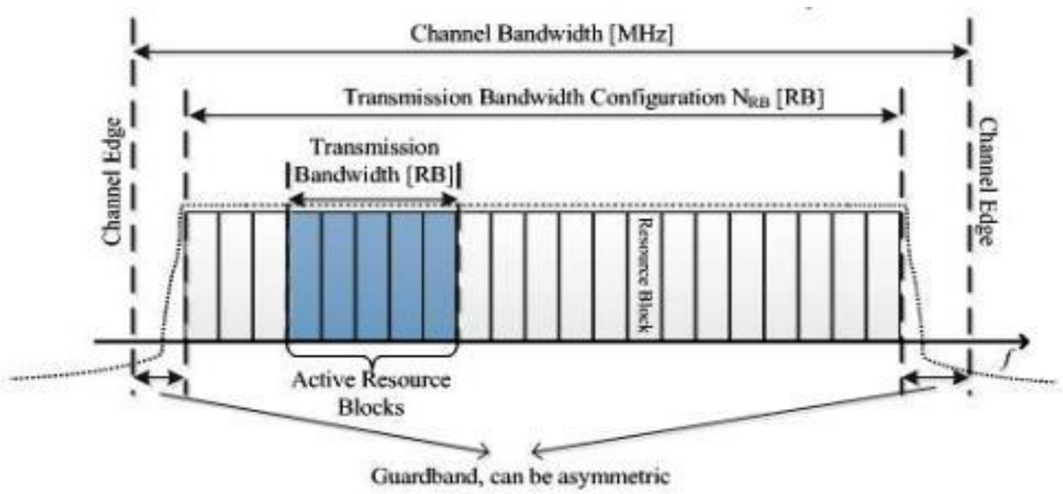


Figure (I.13): 5G-NR channel [9]

I.9 Conclusion:

With the increasing demands of subscribers, 4G will be definitively replaced by 5G using advanced technologies, such as massive MIMO technology, direct communication (Device-to-device), millimeterwave communication, multiple accesses by beam division in massive MIMO technology. The 5G NR numerology provides flexible flexibility with its heterogeneous architecture.

Chapter II

II.1 Introduction:

The objective of the current chapter is to present the massive MIMO (Multiple Input Multiple Output) networks known as Massive MIMO. This emerging technology presents itself as one of the most promising methods for revolutionary fifth generation systems in mobile networks (5G) and the Internet of Things. [13] Technological developments will require powerful and reliable communication systems. This requires a lot of progress on current systems. As part of the possible improvements is the Massive MIMO. Thus, we explore traditional MIMO technology

II.2 Traditional MIMO:

MIMO (Multiple-Input Multiple-Output) technology is a wireless technology that uses multiple transmitters and receivers to transfer more data at the same time. MIMO technology takes advantage of a phenomenon of radio waves called multi-paths, where the transmitted information bounces off walls, ceilings and other objects, reaching the receiving antenna several times from different angles and at slightly different. It exploits multipath behavior by using multiple "smart" transmitters and receivers with an additional "spatial" dimension to dramatically increase performance and range. MIMO allows multiple antennas to send and receive multiple spatial streams at the same time thanks to MIMO, antennas work smarter by allowing them to combine data streams from different paths and at different times to effectively increase the signal capture power of the receiver. Smart antennas use spatial diversity technology, which makes the most of surplus antennas. If there are more antennas than spatial streams, the additional antennas can add diversity to the receivers and increase the range. [14]

II.3 Principles of MIMO:

As explained above, the main source of disturbances to which a signal is subjected during its propagation is the channel. Indeed, because of the multipath propagation phenomena, the signal undergoes fading, frequency or even time offsets. Unlike conventional systems, diversity systems take advantage of these types of propagation to improve system performance. To implement these improvements, MIMO systems use the techniques [15] of:

- ✓ **Space diversity:** Also known as antenna diversity, it means using a several antennas in both transmitter and receiver. [15]

- ✓ **Frequency diversity:** This technique requires sending the same signal on different frequencies. However, you should pay attention to the coherent bandwidth and frequency range due to multipath and the distances to be crossed by the transmission. [15]

- ✓ **Temporal diversity:** When separating the sending of the same signal by the channel coherence time, it is possible to take advantage of temporal diversity. It also depends on the speed of movement of the mobile and the carrier frequency. [15]

II.4 Types of MIMO system:

There are four basic antenna configuration models that include: [4]

- ✓ SISO - Single Input Single Output
- ✓ SIMO - Single entry Multiple exit
- ✓ MISO - Single output with multiple inputs

- ✓ MIMO - Multiple input Multiple output

The term MU-MIMO (Multiuser MIMO) is also used for a multi-user version of the different forms of antenna technology related to one or more inputs and outputs. These are linked to the radio link. In this way, the input is the transmitter when it is transmitted in the link or the signal path and the output is the receiver, it is located at the output of the wireless link. [4]

II.5 Point-to-Point MIMO:

Point-to-Point MIMO emerged in the late 1990s and represents the simplest form of MIMO: a base station equipped with an antenna array serves a terminal equipped with an antenna array; see **Figure (II.1)**. Different terminals are orthogonally multiplexed, for example via a combination of time- and frequency-division multiplexing. In what follows, we summarize some basic facts about Point-to-Point MIMO. In each channel use, a vector is transmitted and a vector is received. In the presence of additive white Gaussian noise at the receiver, Shannon theory yields the following formulas for the link spectral efficiency (in b/s/Hz) [16]:

$$\boldsymbol{C^{uI} = \log_2 \left| I_M + \frac{\rho uI}{K} G G^H \right|} \tag{2.1}$$

$$\boldsymbol{C^{dI} = \log_2 \left| I_K + \frac{\rho dI}{M} G^H G \right| = \log_2 \left| I_M + \frac{\rho dI}{M} G G^H \right|} \tag{2.2}$$

In (2.1) and (2.2), $G$is an (M, K)matrix that represents the frequency response of thechannel between the base station array and the terminal array; ρuI and ρdI are the uplink and downlink signal-to-noise ratios (SNRs), which are proportional to the corresponding total radiated powers; M is the number of base station antennas; and K is the number of terminal antennas. Also, in (a) we used Sylvester's determinant theorem. The normalization by K and M reflects the fact that for constant values of $\rho uI$and ρdI total radiated power is independent of the number of antennas. The spectral efficiency values in (2.1) and (2.2) require the receiver to know Gbut do not require the transmitter to know G. Performance can be improved somewhat if the transmitter also acquires channel state information (CSI). However, this requires special effort and is seldom seen in practice. [16]

In isotropic (rich) scattering propagation environments, well modeled by independent Rayleigh fading, for sufficiently high SNRs (Signal-to-noise ratio), C^{uI} and C^{dI} scale linearly with $\min(M, K)$and logarithmically with the SNR. Hence, in theory, the link spectral efficiency can beincreased by simultaneously using large arrays at the transmitter and the receiver, that is, making M and K large. In practice, however, three factors seriously limit theusefulness of Point-to-Point MIMO, even with large arrays at both ends of the link. First, the terminal equipment is complicated, requiring independent RF chains per antenna as well as the use of advanced digital processing to separate the data streams. Second, more fundamentally, the propagation environment must support min (M, K) independent streams. This is often not the case in practice when compact arrays are used. Line-of-sight (LoS) conditions are

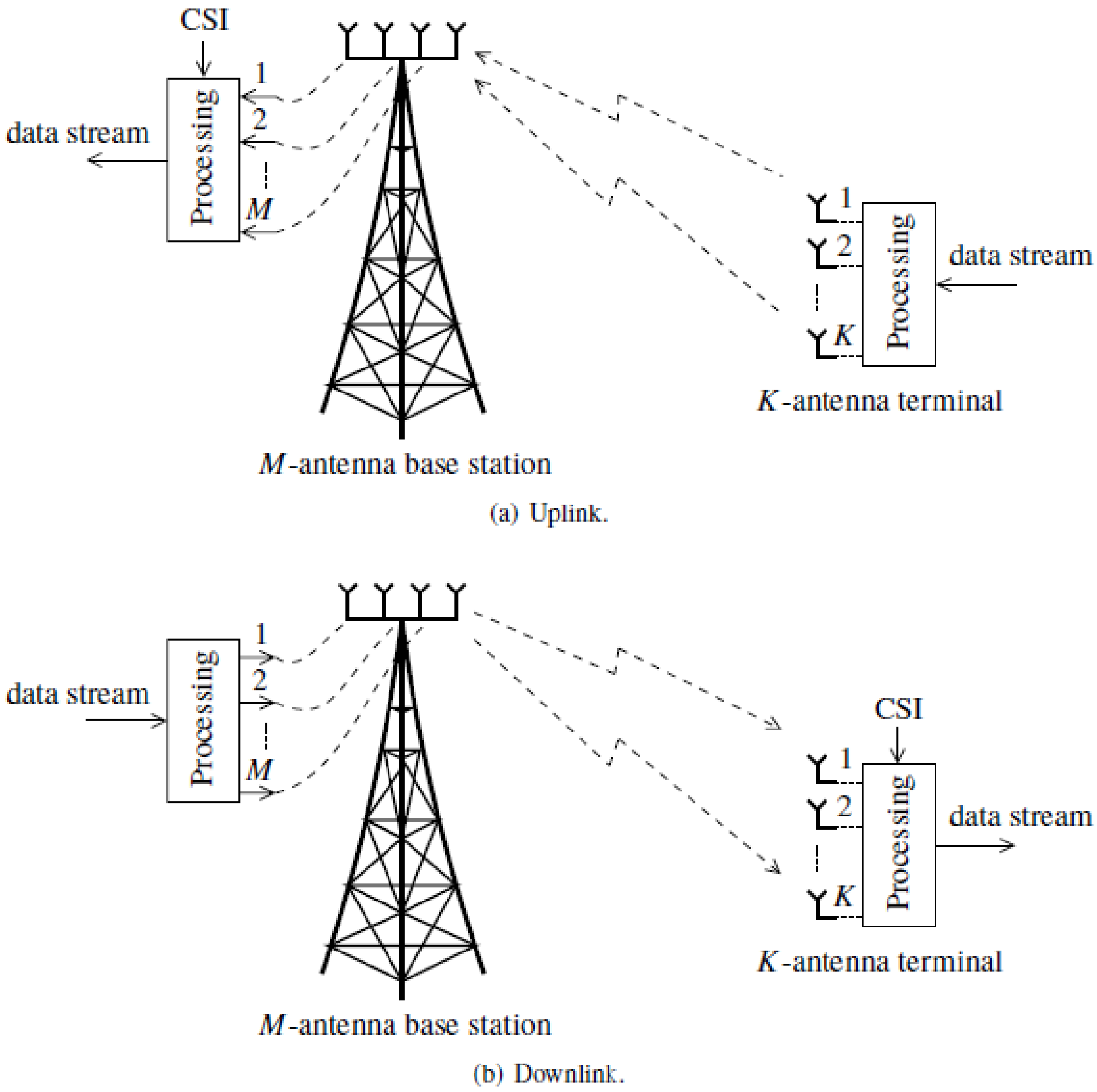


Figure (II.1): Point-to-Point MIMO [16]

Particularly stressing. Third, near the cell edge, where normally a majority of the terminals are located and where SNR is typically low because of high path loss, the spectral efficiency scales slowly with $\min(M, K)$. **Figure (II.2)** illustrates this problem on the downlink for a terminal with $K = 4$ antennas and an SNR of $-3dB$.[16]

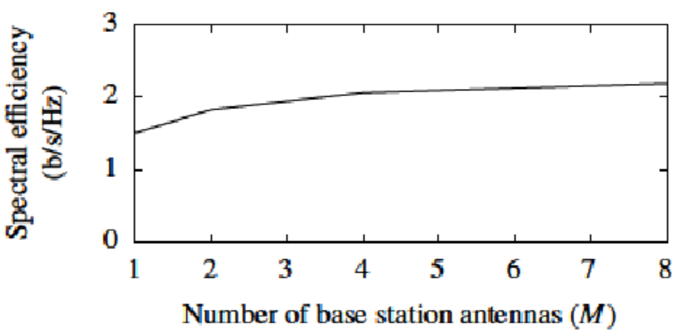


Figure (II.2): Downlink spectral efficiency with Point-to-Point MIMO for a terminal at the cell edge with $K = 4$ antennas, no CSI at the base station, and an SNR of $-3dB$[16]

II.6 Multiuser MIMO (MU-MIMO):

The idea of Multiuser MIMO is for a single base station to serve a multiplicity of terminals using the same time-frequency resources; see **Figure (II.3)**. Effectively, the Multiuser MIMO scenario is obtained from the Point-to-Point MIMO setup by breaking up the K-antenna terminal into multiple autonomous terminals. This section summarizes some basic results of Multiuser MIMO. [16]

Our discussion in this section is confined to that particular form of Multiuser MIMO for which there is a comprehensive Shannon theory which provides the ultimate performance of the system and specifies how this performance may be approached arbitrarily closely. It will be convenient to call this conventional Multiuser MIMO, even though it is doubtful if such a system has ever been reduced to practice. [16]

In **Figure (II.3)** the base station serves K terminals. Let **G** be an $M \times K$matrix corresponding to the frequency response between the base station array and the K [16]

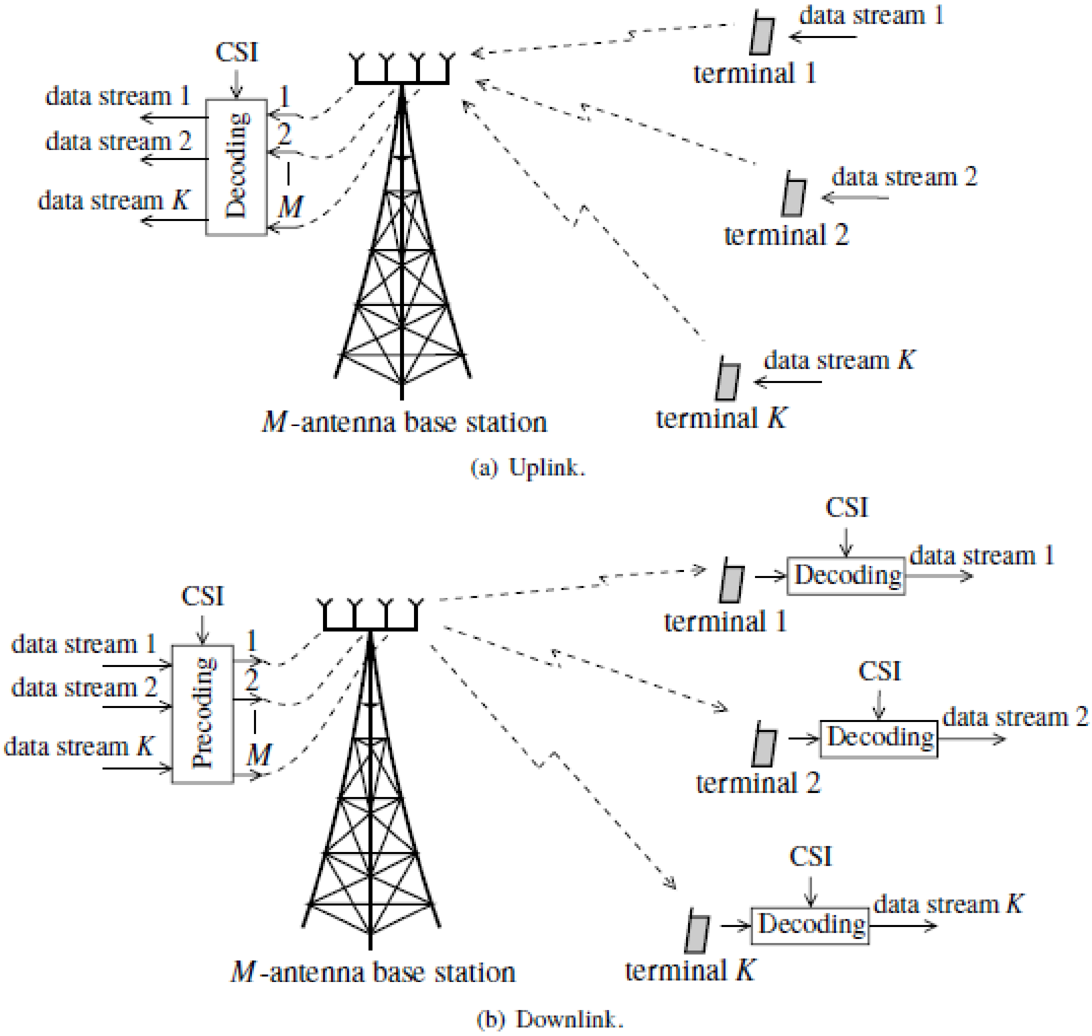


Figure (II.3): Multiuser MIMO. [16]

Terminals. The uplink and downlink sum spectral efficiencies are given by

$$C^{ul} = \log_2 |I_M + p_{ul}\, GG^H| \qquad (2.3)$$

$$C^{dI} = \max_{\substack{V_K \geq 0 \\ \sum_{k=1}^{K} V_K \leq 1}} \log_2 \left| I_M + p_{dI}\, G D_y G^H \right| \quad \textbf{(2.4)}$$

Where $v = [v_1, \dots, v_k]^T$, $p_{uI}$is the uplink SNR per terminal, and p_{dI} is the downlink SNR. (For given p_{uI}, the total uplink power is **K** times greater than for the Point-to- Point MIMO model.) The computation of downlink capacity according to (2.4) requires the solution of a convex optimization problem. The possession of CSI is crucial to both (2.3) and (2.4). On uplink, the base station alone must know the channels, and each terminal has to be told its permissible transmission rate separately. On the downlink, both the base station and the terminals must have CSI.

Note that the terminal antennas in the point-to-point case can cooperate, whereas the terminals in the multiuser case cannot. Quite remarkably, however, the inability of the terminals to cooperate in the multiuser system does not compromise the uplink sum spectral efficiency as seen by comparing (2.1) and (2.3). Note also that the downlink capacity (2.4) may exceed the downlink capacity in (2.2) for Point-to-Point MIMO, because (2.4) assumes that the base station knows G, where as (2.2) does not. Multiuser MIMO has two fundamental advantages over Point-to-Point MIMO. First, itis much less sensitive to assumptions about the propagation environment. For example, LoS conditions are stressing for Point-to-Point MIMO, but not for Multiuser MIMO. Second, Multiuser MIMO requires only single-antenna terminals. Notwithstanding these virtues, two factors seriously limit the practicality of Multiuser MIMO in its originally conceived form. First, to achieve the spectral efficiencies in (2.3) and (2.4) requires complicated signal processing by both the base station and the terminals.

Second, and more seriously, on the downlink both the base station and the terminals must know G, which requires substantial resources to be set aside for transmission of pilots in both directions. For these reasons, the original form of Multiuser MIMO is not scalable either with respect to M or to K. [16]

II.7 MASSIVE MIMO:

II.7.1 What is massive MIMO?

Massive MIMO is a scaled-up version of the conventional small-scale MIMO systems. As shown in **Figure (II.4)**, massive MIMO system is a multiuser communications solution that employs a large number (practically some dozens or hundreds, theoretically up to thousands) of antenna elements to serve simultaneously multiple users with a flexibility to opt what users to schedule for reception at any given time. The most common massive MIMO concept assumes that the user terminals have just a single antenna1 and that the number of antennas at the BS is significantly larger than the number of served users. The introduction of massive MIMO had a tremendous impact on the research and development community during past decade. As a result, many next generation communication technologies, such as 5G below 6 GHz adopted massive MIMO as their key technology. Most of the massive MIMO literature focuses on mobile broadband type-high rate problems with large data packets such that channel estimation and training makes clearly sense. The other application of interest is the massive machine-type communications (mMTC) wherein large number of connected devices are only sporadically active. [18]

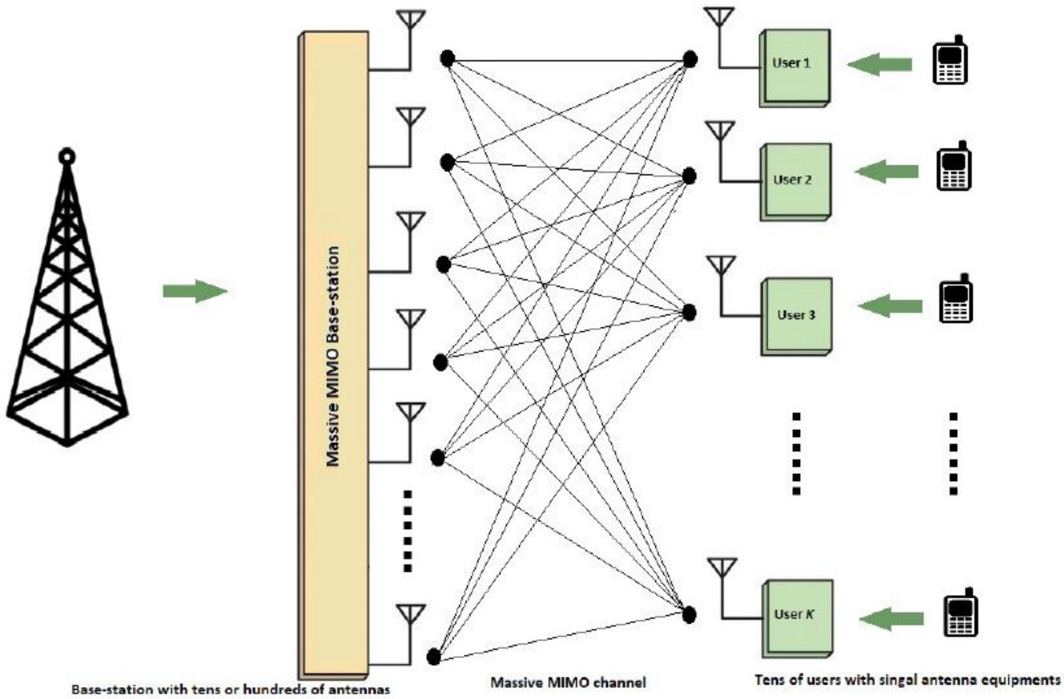


Figure (II.4): Massive MIMO architecture [19]

Some main points of Massive MIMO are [19]:

• **TDD operation (**Time Division Duplexing**):** with FDD (Frequency Division Duplexing), the channel estimation overhead depends on the number of BS antennas, M. By contrast, with TDD; the channel estimation overhead is independent of M. In Massive MIMO, M is large, and hence, TDD operation is preferable. For example, assume that the coherence interval is $T = 200$ symbols (corresponding to a coherence bandwidth of $200\ KHz$ and a coherence time of $1ms$). Then, in FDD systems, the number of BS antennas and the number of users are constrained by $M + K < 200$, while in TDD systems, the constraint on M and K is $2K < 200$. **Figure (II.5)** shows the regions of feasible (M, K) in FDD and TDD systems. We can see that the FDD region is much smaller than the TDD region. With TDD, adding more antennas does not affect the resources needed for the channel estimation. [18]

• **Linear processing:** since the number of BS antennas and the number of users are large, the signal processing at the terminal ends must deal with large dimensional matrices/vectors. Thus, simple signal processing is preferable. In Massive MIMO, linear processing (linear combing schemes in the uplink and linear precoding schemes in the downlink) is nearly optimal. [18]

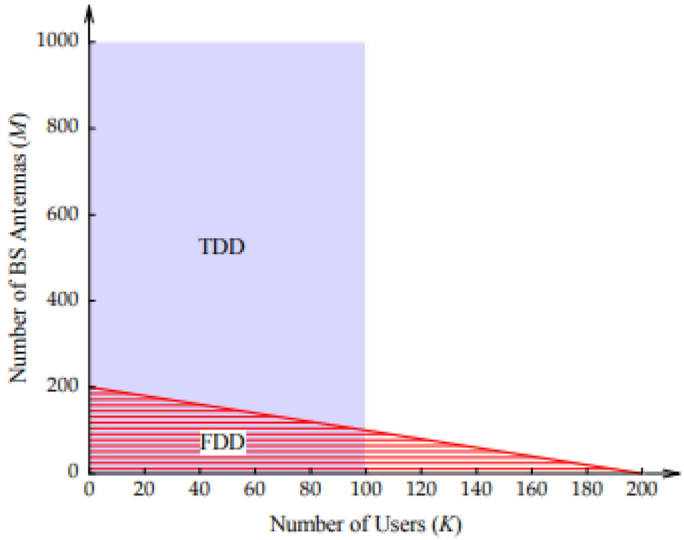


Figure (II.5): The regions of possible (M, K) in TDD and FDD systems, for a coherence interval of 200 symbols. [18]

• **Favorable propagation:** favorable propagation means that the channel matrix between the BS antenna array and the users is well-conditioned. In Massive MIMO, under some conditions, the favorable propagation property holds due to the law of large numbers. [18]

• **A massive BS antenna:** array does not have to be physically large. For example, consider a cylindrical array with 128 antennas, comprising four circles of 16 dual-polarized antenna elements. At 2.6 GHz, the distance between adjacent antennas is about 6 cm, which is half a wavelength, and hence, this array occupies only a physical size of 28cm×29cm. [18]

• **Massive MIMO is scalable:** in Massive MIMO, the BS learns the channels via uplink training, under TDD operation. The time required for channel estimation is independent of the number of BS antennas. Therefore, the number of BS antennas can be made as large as desired with no increase in the channel estimation overhead. Furthermore, the signal processing at each user is very simple and does not depend on other users' existence, i.e., no multiplexing or de-multiplexing signal processing is performed at the users. Adding or dropping some users from service does not sect other users' activities. [18]

• **All the complexity is at the BS**. [18]

II.7.2 Why massive MIMO?

Massive MIMO technology relies on increasing the spatial multiplexing gain and the diversity gain by adding the number of antennas at the BS to serve users with relatively simple processing of signals from all the antennas. The potential benefits of massive MIMO can be summarized as follows [19]:

- **Capacity and link reliability**: Massive MIMO increases the diversity gain, and hence, provides link robustness as it resists fading. It is approved that the capacity increases without a bound as the number of antenna increases, even under a pilot contamination, when multi cell minimum mean square error (MMSE) precoding = combining and spatial channel correlation are used. [26]
- **Spectral efficiency:** Massive MIMO improves the spectral efficiency (SE) of the cellular network by spatial multiplexing of a large number of user equipment’s per cell. Numerous antennas create more spatial data streams, more throughput, more multiplexing gain, and hence achieve high spectral efficiency. It is shown that the overall spectral efficiency in the massive MIMO can be ten times higher than in the conventional MIMO where tens of users will be served simultaneously in the same time-frequency resources. [19]

- **Energy efficiency:** Due to coherent combining, the transmitted power is inversely proportionate to the number of transmit antennas. As the number of transmit antennas increases, the transmit power will be significantly reduced. The power per antenna should be$\propto \frac{\mathbf{1}}{\boldsymbol{n_t}}$, where n_t is the number of antennas. Also, the throughput could be increased by increasing the number of transmit antennas and without increasing the transmit power [20]. Each antenna uses extremely low power, i.e., mill watts. Therefore, the energy efficiency increases and equivalently system reliability.
- **Security enhancement and robustness improvement**: Manmade interference and intentional jamming are serious concerns in modern wireless communication systems. Massive number of antenna terminals leads to a large number of degrees of freedom which can be used to cancel the signals from intentional jammers. In addition, massive MIMO systems are also inherently robust against passive eavesdropping attacks because of beamforming. However, the eavesdropper can take countermeasures by exploiting the high channel correlation in the vicinity of the user or the weakness of channel estimation. [21]
- **Cost efficiency:** Massive MIMO eliminates the need for bulky items such as coaxial cables which used to connect the BS components, and hence reduces the system implementation cost. In addition, massive MIMO uses cheap milliwatts amplifier instead of a multiple expensive high power amplifier [23]. Moreover, it has a potential to reduce the radiated power 1000 times and at the same time drastically maximize the data rates. [22]
- **Signal processing**: A large number of antennas eliminates the interference effects, fast fading, uncorrelated noise and thermal noise, and hence simplifies the signal processing [24] [25]. In addition, it is favorable propagation environment occurs when the channel responses from the base station to user terminals are different (mutually orthogonal, i.e., the inner products are zero). However, non-orthogonal channel vectors lead to advanced signal processing to suppress the interference. [19]

One of the key properties of massive BS antenna arrays is so called channel hardening. It refers to the phenomenon where the massive MIMO channel matrix approaches their expected values, when the number of antennas approaches infinity. In other words, the effective channel approaches deterministic and the off-diagonal components of the Gramian matrix become weaker compared to diagonal terms as the size of the channel gain matrix increases. This property can be exploited in the detection technique and the channel estimation. The simple matched filter (MF) approaches optimality in such a case. However, this is only true with rich scattering and truly large antenna arrays. Therefore, advanced detection techniques are of interest in the practical propagation scenarios and correlated fading channels. [19]

II.7.3 How Massive MIMO Works?

In Massive MIMO, TDD operation is preferable. During a coherence interval, there are three operations: channel estimation (including the uplink training and the downlink training), uplink data transmission, and downlink data transmission. A TDD Massive MIMO protocol is shown in **Figure (II.6)**. [18]

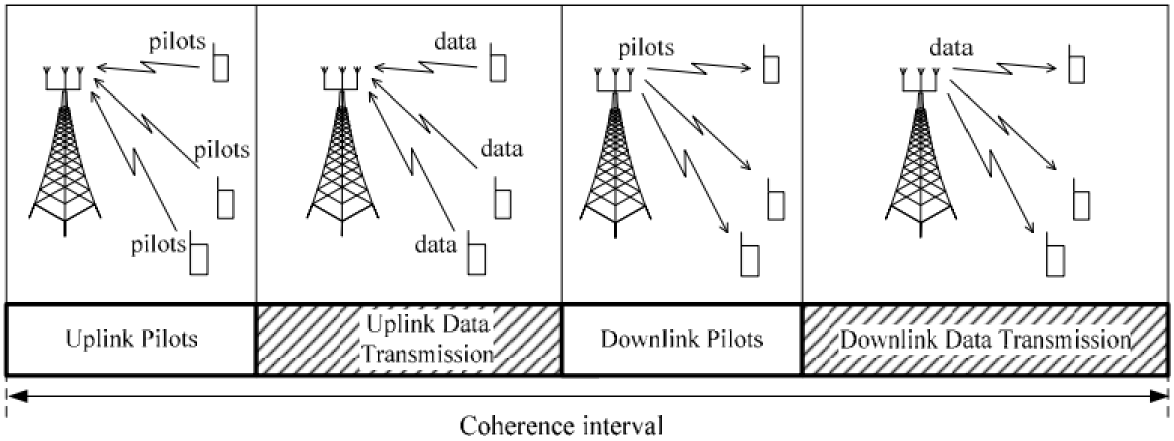


Figure (II.6): Transmission protocol of TDD Massive MIMO. [18]

II.7.3.1 Channel Estimation:

The BS needs CSI to detect the signals transmitted from the users in the uplink, and to precoderthe signals in the downlink. This CSI is obtained through the uplink training. Each user is assigned an orthogonal pilot sequence, and sends this pilot sequence to the BS. The BS knows the pilots sequences transmitted from all users, and then estimates the channels based they received pilot signals. Furthermore, each user may need partial knowledge of CSI to coherently detect the signals transmitted from the BS. This information can be acquired through downlink training or some blind channel estimation algorithm. Since the BS uses linear precoding techniques to beamform the signals to the users, the user needs only the effective channel gain (which is a scalar constant) to detect its desired signals. Therefore, the BS can spend a short time to beamform pilots in the downlink for CSI acquisition at the users. [18]

II.7.3.2 Uplink Data Transmission:

A part of the coherence interval is used for the uplink data transmission. In the uplink, all K users transmit their data to the BS in the same time-frequency resource. The BS then uses the channel estimates together with the linear combining techniques to detect signals transmitted from all users. [18]

II.7.3.3 Downlink Data Transmission:

In the downlink, the BS transmits signals to all K users in the same time-frequency resource. More specifically, the BS uses its channel estimates in combination with the symbols intended for the K users to create M precoded signals which are then fed to M antennas. [19]

II.7.4 Types of Massive MIMO:

II.7.4.1 Single-User MIMO:

Due to the physical inconvenience of the terminals, the number of antennas on the terminal is generally much less than (M). Therefore, the SU-MIMO structures enter in case 1 when a considerable number of antennas are installed at the base station, and therefore enjoy the advantages of channel orthogonalization if the favorable circumstances of channel propagation are valid. However, SU-MIMO channels can be extraordinarily correlated due to the short distance from the terminal side antennas and the viable line-of-sight environment. From an energy efficiency perspective, using a massive array of antennas to service a single or small amount of UT **Figure (II.8)** may also be unwise. Therefore, in this case, the realization of massive MIMO for SU-MIMO can also be limited [19]

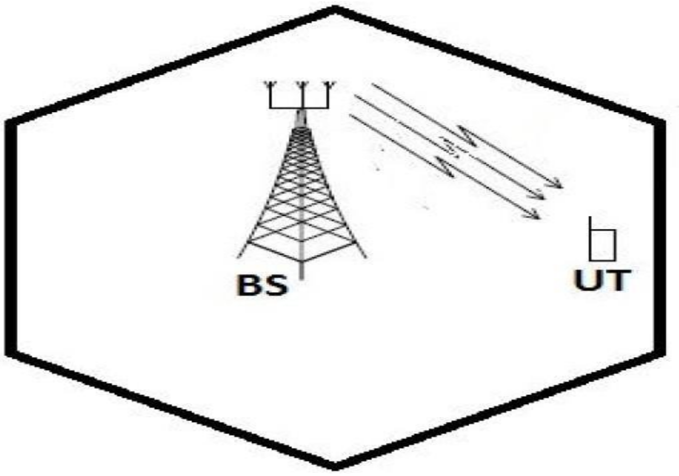


Figure (II.7): Single-User Massive MIMO

$$\boldsymbol{Y} = \sqrt{\boldsymbol{Pu}}\boldsymbol{hs} + \boldsymbol{w} \quad \textbf{(2.5)}$$

Let $(\boldsymbol{Pu})$denote the uplink SNR, (**h**), the channel response vector, (s) is the symbol vector, and (w) is the AWGN noise vector.

II.7.4.2 Multi-User Massive MIMO:

When several terminals are authorized to access an identical time-frequency resource (**Figure (II.8)**), MU-MIMO offers system efficiency higher than that of SU-MIMO. In this section, we consider MU-MIMO single cell systems, in which the base station serves $K\ UTs$, each terminal being equipped with an antenna. The signal received on the base station of an MU-MIMO system on the uplink is: [18]

$$\boldsymbol{y} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{k}} \sqrt{\boldsymbol{Pu}}\boldsymbol{h_k s_k} + \boldsymbol{w} \quad \textbf{(2.6)}$$

$$\boldsymbol{y} = \sqrt{\boldsymbol{Pu}}\boldsymbol{HS} + \boldsymbol{w} \quad \textbf{(2.7)}$$

y is the received signal matrix$M \times 1$,
$\boldsymbol{h_k} \in \boldsymbol{H}$. Or $\boldsymbol{H} = [\boldsymbol{h_1} \dots \boldsymbol{h_k} \dots \boldsymbol{h_K}]$ represents the channel vector between the antennas BS and the $\boldsymbol{k^{th}UT}$,
$\boldsymbol{s_k} \in \boldsymbol{S}$. Or $\boldsymbol{S} = [\boldsymbol{s_1} \dots \boldsymbol{s_k} \dots \boldsymbol{s_K}]$represents the symbol transmitted by the $\boldsymbol{k^{th}UT}$,
w represents the additive white Gaussian noise (AWGN).

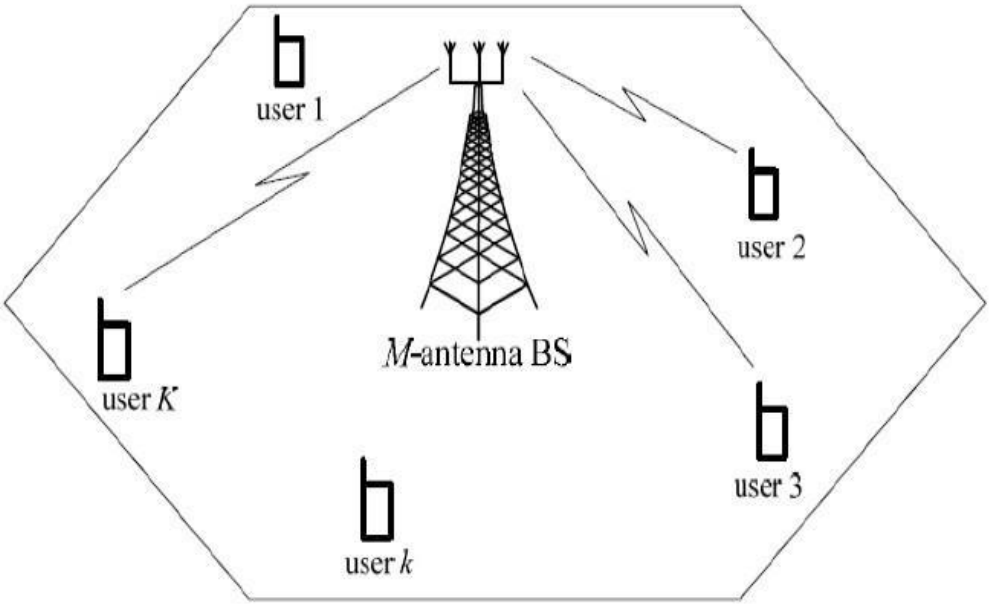


Figure (II.8): Massive MU-MIMO system. M-antennas from BS serves K single antenna UT [18]

When $\boldsymbol{K \geq 2}$, the signal obtained from each terminal interferes with those of the other terminals and we must therefore anticipate that the mutual information of each terminal for MU-MIMO is smaller than that of SU-MIMO with the same power transmitted to each terminal. However, when $\boldsymbol{M \gg K}$, the channel orthogonalization is triggered, so that the signal obtained from each terminal is almost orthogonal, that is to say free from interference in the preferred signaling structure below the conditions favorable channel propagation. In addition, since the terminals are autonomous, the favorable channel propagation condition is generally comfortable since the antennas at the terminals are almost uncorrelated and decoupled. This suggests that Massive MIMO is the wish of the MU-MIMO configuration. [18]

II.7.4.3 MU- Massive MIMO with Multi-Cell scenario:

In this section, we consider restricting non-cooperative cellular multi-user MIMO systems as (M) grows without limit. For a single cell, as well as for multi-cell MIMO, letting (M)increase without limit has the final effect of removing thermal noise and fading it from Rayleigh on a small scale. However, interference from separate cells due to pilot contamination will persist with multiple cells. The idea of pilot pollution is new in a MU-MIMO cellular context and is illustrated in the following **Figure (II.9)**. However, this is usually a problem in the context of CDMA (Code Division Multiple Access), usually under the title of "pilot contamination". The channel estimate calculated using the base station in cell $(\boldsymbol{I})$is contaminated by the pilot transmission from cell $(\boldsymbol{j})$. The base station of cell $(\boldsymbol{I})$will transmit its signal partially alongside the terminals of the adjacent cell. Due to beam formation, the interference to cell $(\boldsymbol{j})$. Does not disappear as a way asymptotic as $\boldsymbol{M} \rightarrow \infty$. A cellular system multiuser MIMO-OFDM with hexagonal cells and NFFT subcarriers is favored. All cells serve $(\boldsymbol{K})$. Independent terminals and have $(\boldsymbol{M})$. Antennas at the base station. Base stations are assumed to be non-cooperative.
The matrix of composite channels $M \times K$ between the $(\boldsymbol{K})\boldsymbol{UTs}$ in cell $(\boldsymbol{I})$ and the BSs in cell $(\boldsymbol{j})$.is denoted $(\boldsymbol{H_{Ij}})$. Based on reciprocity, the downlink channel matrix between the base station of cell $(\boldsymbol{j})$. And the terminals of cell (l) is presented by $(\boldsymbol{H_{Ij}T})$. The signal received at the $\boldsymbol{j^{th}BS}$ will be as follows [18]:

$$\boldsymbol{Y_j = \sum_{I=1}^{L} \sum_{K=1}^{K} \sqrt{Pu}\, h_{jIK} s_{IK} + w} \qquad \textbf{(2.8)}$$

$$Y_j = \sum_{I=1}^{L} \sqrt{Pu}\, h_{jI} s_I + w \tag{2.9}$$

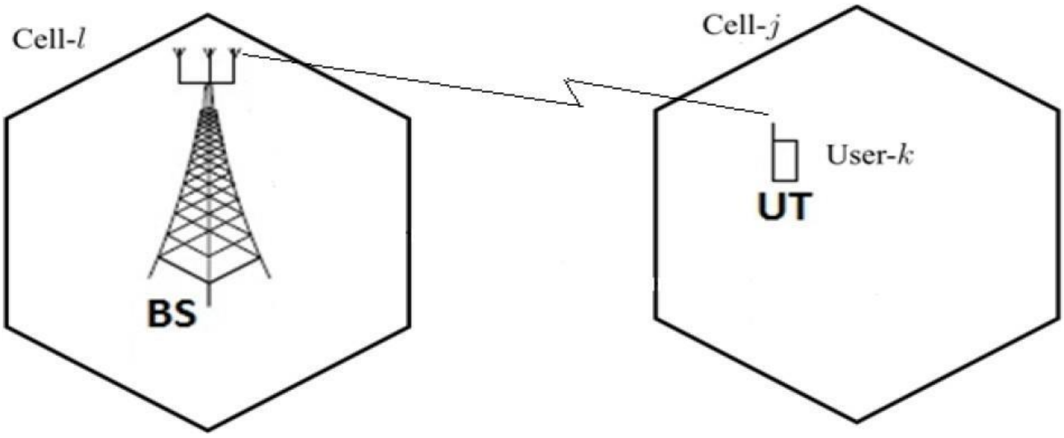


Figure (II.9) : La BS dans la I^{th} cellule et le $K^{th}UT$dans la j^{th} cellule [18]

II.7.5 Massive distributed MIMO:

Distributed Massive MIMO can be treated as a separate case from MU-MIMO in order to additionally provide greater system capacity by using dispersed deployed antennas to transmit and receive signals. One of the mechanisms of disbursed massive MIMO is to enable cooperation between base stations in separate cells, which reduces interference between cells. However, synchronization becomes an essential problem even for antennas distributed in the same base station. In some cases, the massive amount of antennas at the base station can also be positioned in unique places (for example, on top of buildings). In this case, synchronization is a problem, and the inexpensive RF interface can also pose greater problems [17].

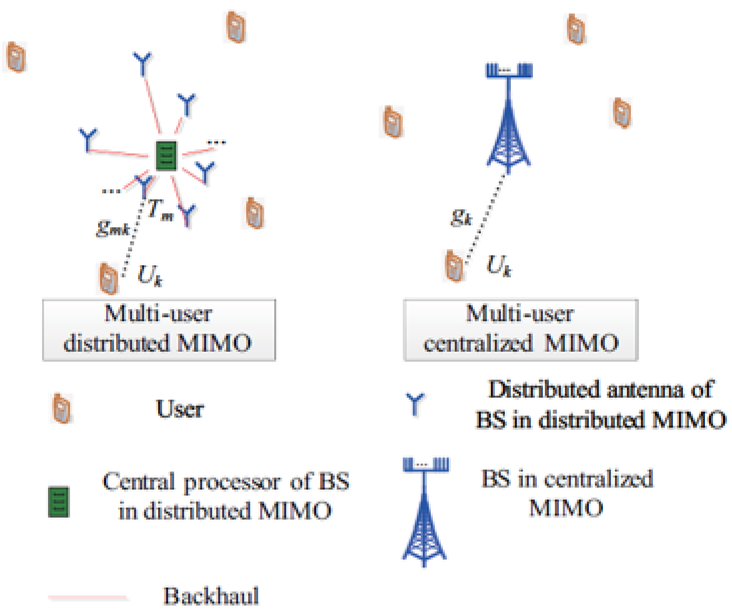


Figure (II.10.a): System models of multi-user distributed MIMO (let side) and multi-user centralized MIMO (right side) [49]

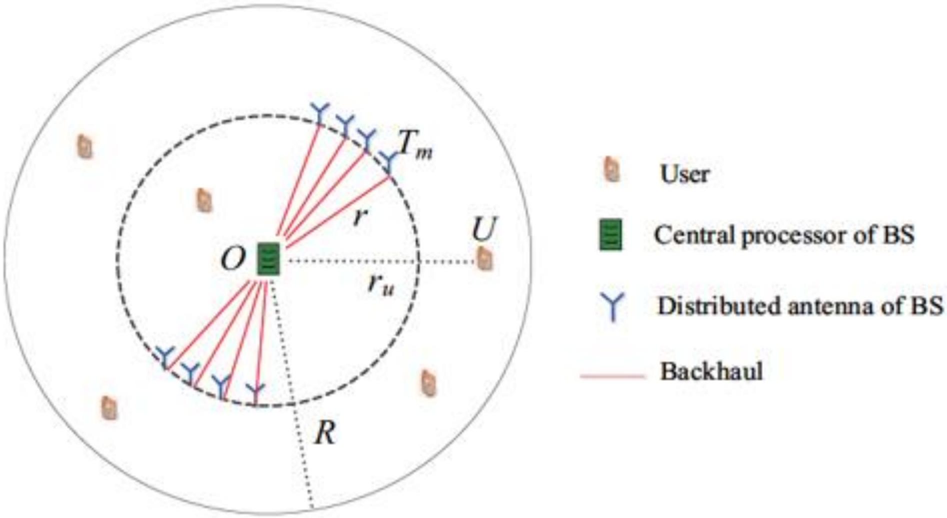


Figure (II.10.b): Distributed massive MIMO system with circular BS antennas [49]

II.7.6 Challenges in Massive MIMO:

Despite the huge advantages of Massive MIMO, many issues still need to be tackled. The main challenges of Massive MIMO are listed as follows. [18]:

II.7.6.1 Pilot Contamination:

In previous sections, we considered single-cell setups. However, practical cellular networks consist of many cells. Owing to the limited availability of frequency spectrum, many cells have to share the same time-frequency resources. Thus, multi cell setups should be considered. In multi cell systems, we cannot 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. This effect, called “pilot contamination”, reduces the system performance. The effect of pilot contamination is major inherent limitation of Massive MIMO. It does not vanish even when the number of BS antennas grows without bound. Considerable sorts have been made to reduce this effect. The eigen value-decomposition-based channel estimation, pilot decontamination, the authors shown that, under certain conditions of the channel covariance, by using a covariance aware pilot assignment scheme among the cells, pilot contamination can be effeciently mitigated. There is much ongoing research on this topic. [18]

II.7.6.2 Unfavorable Propagation:

Massive MIMO works under favorable propagation environments. However, in practice, there may be propagation environments where the channels are not favorable. For example, in propagation environments where the numbers of the scatterers is small compared to the numbers of users, or the channels from different users to the BS share some common scatterers, the channel is not favorable. One possibility to tackle this problem is to distribute the BS antennas over a large area. [18]

II.7.6.3 New standards and designs are required:

It will be very efficient if Massive MIMO can be deployed in current systems such as LTE. However, the LTE standard only allows for up to 8 antenna ports at the BS. Furthermore, LTE uses the channel information that is “assumed”. For example, one option of the downlink in LTE is that the BS transmits the reference signals through several fixed beams. Then the users report back to the BS the strongest beam. The BS will use this beam for the downlink transmission. By contrast, Massive MIMO uses the channel information that is estimated (measured). Therefore, to reduce Massive MIMO to practice, new standards are required. On a different note, with Massive MIMO, a costly 40Watt transceiver should be replaced by a large number of low-power and inexpensive antennas. Related hardware designs should also be considered. This requires a huge effort from both academia and industry. [18]

II.8 Comparison between traditional MIMO and massive MIMO:

Compared to traditional MIMO, the advantages of massive MIMO include [4]:

• Improvement of SE
• Massive amount of degrees of freedom in the space domain
• Good system performance with only a linear (simple) pre-coding scheme, eg Zero forcing, transmission of the maximum ratio, minimum mean square error
• Facilitate the allocation of resources.

II.9 Conclusion:

Massive multiple-input multiple-output (MIMO) technology, where a base station (BS) equipped with very large number of antennas (collocated or distributed) serves many users in the same time-frequency resource, can meet the above requirements, and hence, it is a promising candidate technology for next generations of wireless systems. With massive antenna arrays at the BS, for most propagation environments, the channels become favorable, i.e., the channel vectors between the users and the BS are (nearly) pair wisely orthogonal, and hence, linear processing is nearly optimal. A huge throughput and energy efficiency can be achieved due to the multiplexing gain and the array gain. In particular, with a simple power control scheme, Massive MIMO can offer uniformly good service for all users. [18]

Chapter III

III.1 Introduction:

The massive multiple-input multiple-output (MIMO) system is an important technology in the fifth generation of mobile communication. To get the result of a MIMO system requires some algorithm to approximate the precise result as the computation complexity is too large. There are several methods that have been advanced in the fitting, like zero forcing (ZF) method or minimum mean-square error (MMSE) method. However, in massive MIMO system, these methods require further simplify because of the increasing complexity of matrix inversion. In this chapter, many methods were presented to get the approximation matrix: like MMSE-SIC, ZF-MIC, Gauss-Seidel, Jacobi and Neumann series expansion.

III.2 Detection Algorithms:

In the uplink of the multiuser massive MIMO systems under consideration, the signals or data streams transmitted by the users to the receiver overlap and typically result in multiuser interference at the receiver. This means that the interfering signals cannot be easily demodulated at the receiver unless there is a method to separate them. In order to separate the data streams transmitted by the different users, a designer must resort to detection techniques, which are similar to multiuser detection methods.

III.2.1 Matched Filter detector (MF):

Matched Filter detector (MF) handles the interference from other sub-streams as purely noise by making $\boldsymbol{A} = \boldsymbol{H}$. The estimated received signal using MF is given by:

$$\widehat{\boldsymbol{X}}_{\boldsymbol{MF}} = \boldsymbol{S}(\boldsymbol{H}^{\boldsymbol{H}}\boldsymbol{y}) \quad \textbf{(3.1)}$$

Followed by a slicer **S**, which quantizes each entry to the nearest neighbouring constellation

Which works properly when K(the number of served users) is much smaller than N(the number of receive antennas) and it provides a worse performance compared to more complex detectors. MF, also called the maximum ratio combining (MRC), aims to maximize the received SNR of each stream by neglecting the effect of multiuser interference. If the channel is ill-conditioned, performance is severely degraded for a square MIMO system. [27]

III.2.2 maximum-likelihood detector (ML):

The ML detector has a cost that is exponential in the number of data streams and the modulation order which is too costly for systems with a large number of antennas. Even though the ML solution can be alternatively computed using sphere decoder (SD) algorithms [40]-[39] that are very efficient for MIMO systems with a small number of antennas, the cost of SD algorithms depends on the noise variance, the number of data streams to be detected and the signal constellation, resulting in high computational costs for low SNR values high-order constellations and a large number of data streams.

The high computational cost of the ML detector and the SD algorithms in scenarios with large arrays have motivated the development of numerous alternative strategies for MIMO detection, which are based on the computation of receive filters and interference cancellation strategies. The key advantage of these approaches with receive filters is that the cost is typically not dependent on the modulation, the receive filter is computed only once per data packet and performs detection with the aid of decision thresholds. Algorithms that can compute the parameters of receive filters with low cost are of central importance to massive MIMO systems. In what follows, we will briefly review some relevant suboptimal detectors, which include linear and decision-driven strategies.

III.2.3 Linear ZF detector:

ZF outperforms the MF detector and it aims to maximize the received signal-to-interference ration (SINR). The ZF mechanism is based on inverting the channel matrix $\boldsymbol{H}$and thus, removing the effect of the channel. The equalization matrix of the ZF detector is given by:

$$\boldsymbol{A}_{ZF}^{H} = (\boldsymbol{H}^{H}\boldsymbol{H})^{-1}\boldsymbol{H}^{H} = \boldsymbol{H}^{\dagger} \quad \textbf{(3.2)}$$

Where $\boldsymbol{H}^{\dagger}$ is Pseudo-Inverse of a matrix. The pseudo-inverse is used because $\boldsymbol{H}$is not always a square matrix, i.e. the number of users is not equal to the number of antennas at BS. The estimated signal can be shown as:

$$\widehat{\boldsymbol{X}}_{\mathrm{ZF}}=\mathbf{S}\ (\boldsymbol{A}_{ZF}^{H}\boldsymbol{y}) \quad \textbf{(3.3)}$$

It is clear that the ZF detector neglects the effect of noise and it works properly in interference-limited scenarios in expenses of higher computational complexity. However, the ZF detector and the MF may produce a noise enhancement in case of a small-valued coefficient channel. Therefore, MMSE detector is proposed to take the effect of noise in the equalization process. [19]

III.2.4 Linear MMSE detector:

The main idea of the MMSE detector is to minimize the mean-square error (MSE) between the transmitted x and the estimated signal $H^{H}y$ as given by:

$$\boldsymbol{A}_{MMSE}^{H} = \underset{\mathbf{H}\in\boldsymbol{C}^{N\times K}}{\mathbf{arg}}\ \boldsymbol{min}\, \mathbb{E}\|\mathbf{x} - \mathbf{H}^{\mathbf{H}}\mathbf{y}\|^{2} \quad \textbf{(3.4)}$$

Where $\mathbb{E}$ is the expectation operator. MMSE detector takes the noise effect into consideration as

$$\boldsymbol{A}_{MMSE}^{H} = \left((\boldsymbol{H}^{H}\boldsymbol{H} + \frac{\boldsymbol{K}}{\boldsymbol{SNR}}\boldsymbol{I})^{-1}\right)\boldsymbol{H}^{H} \quad \textbf{(3.5)}$$

Where $\boldsymbol{I}$ is the identity matrix. The output of the MMSE detector can be obtained by:

$$\widehat{\boldsymbol{X}}_{MMSE} = \boldsymbol{S}(\boldsymbol{A}_{MMSE}^{H}\boldsymbol{y}) \quad \textbf{(3.6)}$$

Unlike the ZF detector in (3.3), the MMSE in (3.6) depends on a reduced noise enhancement and it requires knowledge of the SNR. Therefore, the MMSE detector is capable of achieving a significantly better performance than the ZF detector when the noise power is large. [19]

III.3 Massive MIMO Detectors based on approximated matrix inverse:

With a large number of transmit antennas, the channel hardening phenomenon can be exploited to cancel the characteristics of a small-scale fading [39] and it becomes dominant when K is much lower than N. This can be seen as a diagonalization of the entries in the Gram matrix or Gramian G = HHH, where the non-diagonal components tend to zero

And diagonal terms become closer to N [28][29]. As shown in (3.2), a matrix inversion of the Gramian matrix is required to equalize the received signal. It exhibits high computational complexity being one of the most complex operations in the linear and simple non-linear MIMO detectors. For the massive MIMO system, this problem becomes more severe as the dimension of the Gramian G increases

[30]. Several methods have been proposed to reduce complexity by approximating the inverse of a matrix, rather than computing it [31]. Besides the cost of a matrix inversion, a challenge in matrix inversion lies on when the channel matrix is nearly singular and the system becomes ill-conditioned. In this case, the matrix inversion will not equalize the received signal [32], [33]. In order to overcome the inherent noise enhancement, modified detectors with approximate matrix inversion methods will be an essential. Therefore, detectors based on approximate matrix inversion will be presented and discussed below.

III.3.1 System Model:

We consider an uncoded MIMO system where the BS is furnished with N antennas to serve K single antenna users simultaneously in a single cell where$K \ll N$. The vector $\boldsymbol{x}$presents the transmitted data by all users and the symbol vector $\boldsymbol{y}$presents received data at the BS where $\boldsymbol{x}$and $\boldsymbol{y}$ are Kand N, respectively. The received vector $\boldsymbol{y}$is usually corrupted by channel effects and the noise $\boldsymbol{w}$. The channel matrix$\boldsymbol{H}$ entries are independent and identically distributed (i.i.d) with zero mean and unit variance. The MIMO model is given as [34]:

$$\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{w} \tag{3.7}$$

The equalization matrix $\boldsymbol{A}$ plays a role in estimating the signal as:

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^{-1} \boldsymbol{y}_{\boldsymbol{MF}} \tag{3.8}$$

With: $\boldsymbol{A} = \boldsymbol{G} + \boldsymbol{\sigma}^2 \boldsymbol{I}_{\boldsymbol{K}}$ and $\boldsymbol{y}_{\boldsymbol{MF}} = \boldsymbol{H}^{\boldsymbol{H}} \boldsymbol{y}$ **(3.9)**

Where $\boldsymbol{\sigma}^2$ is the noise variance and $\boldsymbol{I}_{\boldsymbol{K}}$ is the K ×K identity matrix. The Gram matrix (**G**) is $\boldsymbol{H}^{\boldsymbol{H}}\boldsymbol{H}$. However, a direct computation of $\mathbf{A}^{-1}$ requires $O(K^3)$. However, literature is rich with methods to avoid the direct and exact matrix inversion which relieves a burden of high computational complexity. [34]

III.3.2 Neumann Series (NS) method:

The Neumann Series (NS) method approximates iteratively the matrix inversion where the problem is represented as a sum of infinite number of terms. However, the computational complexity increases as the number of iterations (n) and terms increase. In such approximation method, the equalization matrix is decomposed into a diagonal matrix (**D**) and a non-diagonal matrix **E**, where $\mathbf{A}^{-1}$ can be iteratively approximated and refined as:

$$\boldsymbol{A}^{-1}_{(\boldsymbol{n}+\boldsymbol{1})} = \sum_{\boldsymbol{n}=\boldsymbol{0}}^{\infty} (\boldsymbol{I} - \boldsymbol{A}^{-1}_{\boldsymbol{n}} \boldsymbol{A})^{\boldsymbol{n}} \boldsymbol{A}^{-1}_{\boldsymbol{n}} \tag{3.10}$$

The initial vector of the matrix inversion $(\boldsymbol{A}^{-1}_{(\boldsymbol{0})})$ is usually selected as D^{-1} and refined iteratively [35].However, it converges to A^{-1} if$\left\| \mathrm{I} - \boldsymbol{A}\boldsymbol{A}^{-1}_{(\boldsymbol{0})} \right\| < 1$. After that, the estimated signal $\hat{\boldsymbol{x}}$ is written as

$$\hat{\boldsymbol{x}} = A^{-1} y_{MF}$$

III.3.3 Gauss–Seidel method:

Gauss–Seidel (GS) method is an efficient iterative method and it is a special scenario of the successive over relaxation (SOR) method.
Signal estimation using the SOR depends on the lower triangular matrix (**L**) and upper triangular matrix (**U**) as:

$$\hat{\boldsymbol{x}}_{(\boldsymbol{n}-\mathbf{1})} = \left(\boldsymbol{D} - \frac{\mathbf{1}}{\boldsymbol{w}}\boldsymbol{L}\right)^{-\mathbf{1}} \left(\boldsymbol{y}_{\mathbf{MF}} + \left(\left(\mathbf{1} - \frac{\mathbf{1}}{\mathbf{w}}\right)\mathbf{D} + \frac{\mathbf{1}}{\mathbf{w}}\mathbf{U}\right)\hat{\boldsymbol{x}}_{(\boldsymbol{n})}\right) \quad \textbf{(3.11)}$$

Where $\boldsymbol{w}$ is the relaxation parameter, the lower triangular matrix. When$\boldsymbol{w} = \mathbf{1}$ in Equation (3.11), the GS iterative method is obtained and the estimated signal can be written as:

$$\hat{\boldsymbol{x}}_{(\boldsymbol{n}-\mathbf{1})}=(\boldsymbol{D} - \boldsymbol{L})^{-1}\left(\boldsymbol{y}_{\boldsymbol{MF}} + \boldsymbol{U}\hat{\boldsymbol{x}}_{(\boldsymbol{n})}\right) \quad \textbf{(3.12)}$$

Generally, the GS method has a fast convergence rate. The initial estimation is sat as $\hat{\boldsymbol{x}}_{\mathbf{(0)}}= D^{-1}y_{MF}$ and can be refined iteratively. [34]

III.3.4 Jacobi method:

The Jacobi method (JA) is another iterative method where estimation of the signal is written as:

$$\hat{\boldsymbol{x}}_{(\boldsymbol{n}+\mathbf{1})} = \boldsymbol{D}^{-\mathbf{1}}\left(\boldsymbol{y}_{\boldsymbol{MF}} + (\boldsymbol{D} - \boldsymbol{A})\hat{\boldsymbol{x}}_{(\boldsymbol{n})}\right) \quad \textbf{(3.13)}$$

The rate of convergence of the JA method is slower than the GS method. However, computational complexity of the JA method is lower than the GS method and it is easy to implement in parallel manner. [36]

III.3.5 Stair Matrix method:

To begin with, we have the following definitions.
Definition 1 [34]: In an N×N matrix $\boldsymbol{A}$, if it-s entry $\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{n})}=\boldsymbol{e}_{\boldsymbol{m}}^{\boldsymbol{H}}\boldsymbol{A}\boldsymbol{e}_{\boldsymbol{n}}$,*m*, *n* =1,2,⋯, N, satis-fies$A_{(m,n)} = 0$where $n \notin \{m-1, m, m+1\}$, wethen call it as a tridiagonal matrix, denoted by
$\boldsymbol{A} = \boldsymbol{tridiag}\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}-\mathbf{1})}, \boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m})}, \boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}+\mathbf{1})}$
A typical tridiagonal matrix *A*can be given as

$$\begin{bmatrix} \boldsymbol{A}_{(\mathbf{1}.\mathbf{1})}\boldsymbol{A}_{(\mathbf{1}.\mathbf{2})} \\ \boldsymbol{A}_{(\mathbf{2}.\mathbf{1})}\boldsymbol{A}_{(\mathbf{2}.\mathbf{2})}\boldsymbol{A}_{(\mathbf{2}.\mathbf{3})} \\ \ddots \quad \ddots \quad \ddots \\ \boldsymbol{A}_{(\boldsymbol{N}-\mathbf{1},\boldsymbol{N}-\mathbf{2})}\boldsymbol{A}_{(\boldsymbol{N}-\mathbf{1},\boldsymbol{N}-\mathbf{1})}\boldsymbol{A}_{(\boldsymbol{N}-\mathbf{1},\boldsymbol{N})} \\ \boldsymbol{A}_{(\boldsymbol{N},\boldsymbol{N}-\mathbf{1})}\boldsymbol{A}_{(\boldsymbol{N},\boldsymbol{N})} \end{bmatrix}$$

Definition 2[36]: If a tridiagonal matrix satisfies one of the following conditions:
$\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}-\mathbf{1})} = \mathbf{0}$, $\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}+\mathbf{1})} = \mathbf{0}$. where *m*=2*k*-1,*k*=1,2,…….,[(N+1)/2]. (I)
Alternatively, the non-diagonal elements in the odd rows of tridiagonal matrix are zeros;
$\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}-\mathbf{1})} = \mathbf{0}$, $\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}+\mathbf{1})} = \mathbf{0}$, where *m*=2*k*, *k*=1, 2, [N/2]. (II)
In other words, the non-diagonal elements in the even rows of tridiagonal matrix are zeros;
We then call it as a stair matrix, denoted by ($\mathbf{A} = \mathbf{stair}\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}-\mathbf{1})}$, $\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m})}$, $\boldsymbol{A}_{(\boldsymbol{m},\boldsymbol{m}+\mathbf{1})}$).

In accordance, a stair matrix is of type I if the condition (I) is satisfied and is of type II if the condition (II) is satisfied.
For example: a **5×5** stair matrix has the following forms [36]:

$$A=\begin{bmatrix} \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & 0 & 0 \\ 0 & 0 & \times & 0 & 0 \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix} \qquad \text{I}$$

$$A=\begin{bmatrix} \times & \times & 0 & 0 & 0 \\ 0 & \times & 0 & 0 & 0 \\ 0 & \times & \times & \times & 0 \\ 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & \times & \times \end{bmatrix} \qquad \text{II}$$

III.4 Estimation of Receive Filter Parameters based on reduced rank algorithm:

An alternative to channel estimation techniques is the direct computation of the receive filters using LS techniques or adaptive algorithms. In this subsection, we consider the estimation of the receive filters for multiuser Massive MIMO systems. The receive filter estimation problem corresponds to solving the LS optimization problem described by [44]:

$$\boldsymbol{w}_{k,o}[\boldsymbol{i}] = \underset{\boldsymbol{w}_k[\boldsymbol{i}]}{\textbf{arg min}} \sum_{l=1}^{i} \boldsymbol{\lambda}^{i-l} \left| \boldsymbol{s}_k[\boldsymbol{l}] - \boldsymbol{w}_k^H[\boldsymbol{i}]\boldsymbol{r}[\boldsymbol{i}] \right|^2 \tag{3.16}$$

Where the $NA \times 1$ vector $\boldsymbol{w}_k$ contains the parameters of the receive filters for the $\boldsymbol{k}^{\text{th}}$ data stream, the symbol $\boldsymbol{s}_k[\boldsymbol{i}]$ contains the symbols of the k^{th} data stream. Similarly to channel estimation, it is common to use known pilot symbols $\boldsymbol{s}_k[\boldsymbol{i}]$ in the beginning of the transmission for estimation of the receiver filters. This problem can be solved by computing the gradient terms of (3.16), equating them to a null vector and manipulating the terms which yields the LS estimate [43]

$$\boldsymbol{w}_{k,o}[\boldsymbol{i}] = \boldsymbol{R}_r^{-1}[\boldsymbol{i}]\boldsymbol{p}_k[\boldsymbol{i}] \tag{3.17}$$

Where $\boldsymbol{R}_r[\boldsymbol{i}] = \sum_{l=1}^{i} \boldsymbol{\lambda}^{i-l}\boldsymbol{r}[\boldsymbol{i}]\,\boldsymbol{r}^H[\boldsymbol{i}]$ is the auto-correlation matrix of the received data and $\boldsymbol{p}_k[\boldsymbol{i}] = \sum_{l=1}^{i} \boldsymbol{r}[\boldsymbol{i}]\,\boldsymbol{s}_k^H[\boldsymbol{i}]$ is a $NA \times 1$ vector with cross-correlations between the pilots and the received data $\boldsymbol{r}[\boldsymbol{i}]$. When the channel is static over the duration of the transmission, it is common to set the forgetting factor $\boldsymbol{\lambda}$ to one. Conversely, when the channel is time-varying one needs to set $\boldsymbol{\lambda}$ to a value that corresponds to the coherence time of the channel in order to track the channel variations. In these situations, a designer can also compute the parameters recursively, thereby taking advantage of the previously computed LS estimates and leading to the RLS algorithm [45] given by:

$$\boldsymbol{k}[\boldsymbol{i}] = \frac{\boldsymbol{\lambda}^{-1}\boldsymbol{P}[\boldsymbol{i}-\boldsymbol{1}]\boldsymbol{r}[\boldsymbol{i}]}{\boldsymbol{1}+\boldsymbol{\lambda}^{-1}\boldsymbol{r}^H[\boldsymbol{i}]\boldsymbol{P}[\boldsymbol{i}-\boldsymbol{1}]\boldsymbol{r}[\boldsymbol{i}]} \tag{3.18.a}$$

$$\boldsymbol{P}[\boldsymbol{i}] = \boldsymbol{\lambda}^{-1}\boldsymbol{P}[\boldsymbol{i}-\boldsymbol{1}] - \boldsymbol{\lambda}^{-1}\boldsymbol{k}[\boldsymbol{i}]\boldsymbol{r}^H[\boldsymbol{i}]\boldsymbol{P}[\boldsymbol{i}-\boldsymbol{1}] \tag{3.18.b}$$

$$\boldsymbol{w}_k[\boldsymbol{i}] = \boldsymbol{w}_k[\boldsymbol{i}-\boldsymbol{1}] - \boldsymbol{k}[\boldsymbol{i}]\boldsymbol{e}_{k,a}^*[\boldsymbol{i}] \tag{3.18.c}$$

Where: $\boldsymbol{e}_{k,a}[\boldsymbol{i}] = \boldsymbol{s}_k[\boldsymbol{i}] - \boldsymbol{w}_k^H[\boldsymbol{i}-\boldsymbol{1}]\boldsymbol{r}[\boldsymbol{i}]$ is the *a priori* error signal for the k^{th} data stream. Several other variants of the RLS algorithm could be used to compute the parameters of the receive filters [46]. The computational cost of this RLS algorithm for all data streams corresponds to $KN_U(3N_A^2 + 4NA + 1)$

multiplications and $KN_U(3N_A^2 + 2NA - 1) + 2NAKN_U$ additions A reduced complexity alternative to the RLS algorithms is to employ the LMS algorithm to estimate the parameters of the receive filters. Consider the mean-square error (MSE)-based optimization problem:

$$\boldsymbol{w}_{k,o}[\boldsymbol{i}] = \underset{\boldsymbol{w}_k[i]}{\arg\min} \, \boldsymbol{E}\left[\left|\boldsymbol{s}_{\boldsymbol{k}}[\boldsymbol{i}] - \boldsymbol{w}_{\boldsymbol{k}}^{\boldsymbol{H}}[\boldsymbol{i}]\boldsymbol{r}[\boldsymbol{i}]\right|^2\right] \quad \textbf{(3.19)}$$

Similarly to the case of channel estimation, this problem can be solved by computing the instantaneous gradient terms of (3.19), using a gradient descent rule and manipulating the terms which results in the LMS estimation algorithm given by:

$$\boldsymbol{w}_{\boldsymbol{k}}[\boldsymbol{i}+\mathbf{1}] = \widehat{\boldsymbol{w}}_{\boldsymbol{k}}[\boldsymbol{i}] + \mu \boldsymbol{e}_{\boldsymbol{k}}^*[\boldsymbol{i}]\ \mathbf{r}[\boldsymbol{i}] \quad \textbf{(3.20)}$$

where the error signal for the k th data stream is $\boldsymbol{e}_{\boldsymbol{k}}[\boldsymbol{i}] = \boldsymbol{s}_{\boldsymbol{k}}[\boldsymbol{i}] - \boldsymbol{w}_{\boldsymbol{k}}^{\boldsymbol{H}}[\boldsymbol{i}]\boldsymbol{r}[\boldsymbol{i}]$ and the step size μshould be chosen between 0 and$2/tr[\boldsymbol{R}]$ [45]. The cost of the LMS estimation algorithm in this scheme is $KN_U(N_A+ 1)$ multiplications and KN_UN_Aadditions [44].

In parameter estimation problems with a large number of parameters such as those found in massive MIMO systems, an effective technique is to employ reduced-rank algorithms which perform dimensionality reduction followed by parameter estimation with a reduced number of parameters. Consider the mean-square error (MSE)-based optimization problem:

$$w_{\boldsymbol{k},\mathbf{0}}[\boldsymbol{i}] = \ \boldsymbol{arg\ min}_{w_{k[i]}}\ \boldsymbol{E}\ [\left|\boldsymbol{s}_{\boldsymbol{k}}[\boldsymbol{i}] - \ \boldsymbol{w}_{\boldsymbol{k}}^{\boldsymbol{H}}[\boldsymbol{i}]\boldsymbol{r}[\boldsymbol{i}]\right|^{\mathbf{2}}] \quad \textbf{(3.21)}$$

where $\boldsymbol{T}_{\boldsymbol{D},\boldsymbol{k}}[\boldsymbol{i}]$is an $NA \times D$ matrix that performs dimensionality reduction and$\boldsymbol{w}_{\boldsymbol{k}}[\boldsymbol{i}]$ is a $D \times 1$ parameter vector. Given $\boldsymbol{T}_{\boldsymbol{D},\boldsymbol{k}}[\boldsymbol{i}]$, a generic reduced-rank RLS algorithm [43]
with D-dimensional quantities can be obtained from (3.18.a)-(3.18.c) by substituting the NA×1 received vector $\boldsymbol{r}[\boldsymbol{i}]$by the reduceddimension $D \times 1$ vector

$$\bar{\boldsymbol{r}}[\mathbf{i}] = \mathbf{T}_{\boldsymbol{D}}^{\boldsymbol{H}}\ ,\boldsymbol{k}[\boldsymbol{i}]\boldsymbol{r}[\boldsymbol{i}] \quad \textbf{(3.22)}$$

A central design problem is how to compute the dimensionality reduction matrix $\boldsymbol{T}_{\boldsymbol{D},\boldsymbol{k}}[\boldsymbol{i}]$and several techniques have been considered in the literature, namely:

- Principal components (PC)
- Krylov subspace techniques (KS) or (KST)

III.4.1 Principal Component (PC):

A. What is Principal Component Analysis?

Principal Component Analysis, or PCA, is a dimensionality-reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set [37]

Reducing the number of variables of a data set naturally comes at the expense of accuracy, but the trick in dimensionality reduction is to trade a little accuracy for simplicity. Because smaller data sets are easier to explore and visualize and make analyzing data much easier and faster in the case of machine learning algorithms without extraneous variables to process [37]

So, to sum up, the idea of PCA is simple which can reduce the number of variables of a data set, while preserving as much information as possible.

B. Step by step explanation of PCA:

✓ Step 1: Standardization

The aim of this step is to standardize the range of the continuous initial variables so that each one of them contributes equally to the analysis. [37]

More specifically, the reason why it is critical to perform standardization prior to PCA, is that the latter is quite sensitive regarding the variances of the initial variables. That is, if there are large differences between the ranges of initial variables, those variables with larger ranges will dominate over those with small ranges (For example, a variable that ranges between 0 and 100 will dominate over a variable that ranges between 0 and 1), which will lead to biased results. So, transforming the data to comparable scales can prevent this problem. [37]

Mathematically, this can be done by subtracting the mean and dividing by the standard deviation for each value of each variable. [37]

$$\boldsymbol{z = \frac{value - mean}{standard\ deviation}}$$

Once the standardization is done, all the variables will be transformed to the same scale.

✓ Step 2: Covariance Matrix Computation

The aim of this step is to understand how the variables of the input data set are varying from the mean with respect to each other, or in other words, to see if there is any relationship between them. Because sometimes, variables are highly correlated in such a way that they contain redundant information. So, in order to identify these correlations, we compute the covariance matrix. [37]

The covariance matrix is a $\boldsymbol{p} \times \boldsymbol{p}$ symmetric matrix (where p is the number of dimensions) that has as entries the covariance's associated with all possible pairs of the initial variables. For example, for a 3-dimensional data set with **3** variables $\boldsymbol{x}$, $\boldsymbol{y}$, and z, the covariance matrix is a **3×3** matrix of this from:

$$\boldsymbol{\begin{bmatrix} Cov(x,x) & Cov(x,y) & Cov(x,z) \\ Cov(y,x) & Cov(y,y) & Cov(y,z) \\ Cov(z,x) & Cov(z,y) & Cov(z,z) \end{bmatrix}}$$

Covariance Matrix for 3-Dimensional Data

Since the covariance of a variable with itself is its variance (Cov(a,a)=Var(a)), in the main diagonal (Top left to bottom right) we actually have the variances of each initial variable. And since the covariance is commutative (Cov(a,b)=Cov(b,a)), the entries of the covariance matrix are symmetric with respect to the main diagonal, which means that the upper and the lower triangular portions are equal. [37]

✓ Step 3: Compute the Eigenvectors and Eigen values of the Covariance matrix to identify the Principal Components

Eigenvectors and eigen values are the linear algebra concepts that we need to compute from the covariance matrix in order to determine the Principal Components (PC) of the data. Before getting to the explanation of these concepts, let's first understand what we mean by principal components.
Principal components are new variables that are constructed as linear combinations or mixtures of the initial variables. These combinations are done in such a way that the new variables (i.e., principal components) are uncorrelated and most of the information within the initial variables is squeezed or compressed into the first components. So, the idea is 10-dimensional data gives you 10 principal components, but PCA tries to put maximum possible information in the first component, then maximum remaining information in the second and so on, until having something like shown in the screen plot below. [37]

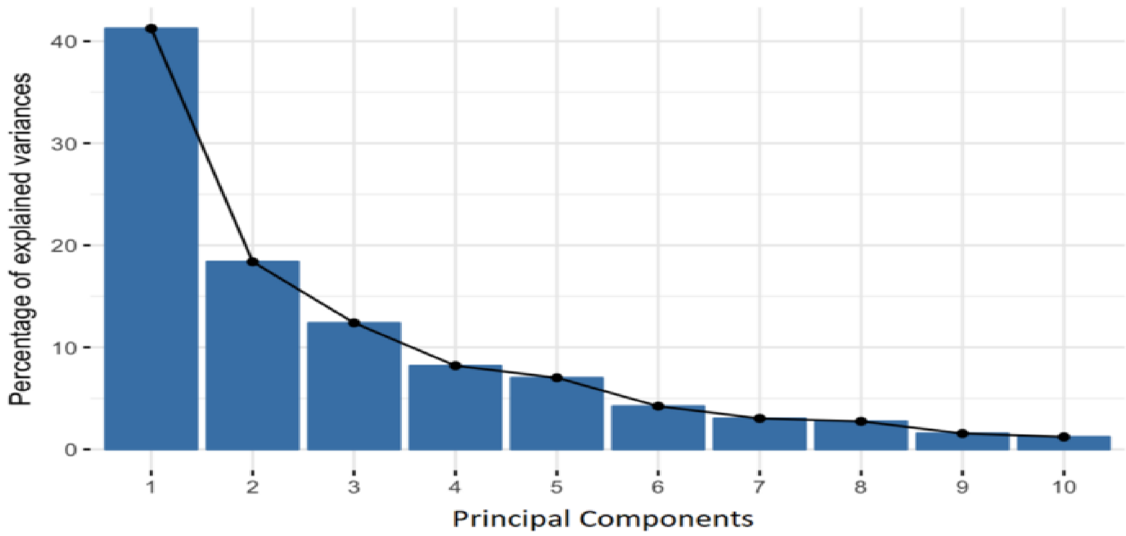


Figure (III.1): Percentage of Variance (Information) for each by PC [37]

Organizing information in principal components this way, will allow to reduce dimensionality without losing much information, and this by discarding the components with low information and considering the remaining components as your new variables. [37]

An important thing to realize here is that, the principal components are less interpretable and don't have any real meaning since they are constructed as linear combinations of the initial variables.

Geometrically speaking, principal components represent the directions of the data that explain a maximal amount of variance, that is to say, the lines that capture most information of the data. The relationship between variance and information here, is that, the larger the variance carried by a line, the larger the dispersion of the data points along it, and the larger the dispersion along a line, the more the information it has. To put all this simply, just think of principal components as new axes that provide the best angle to see and evaluate the data, so that the differences between the observations are better visible. [37]

Let's suppose as an example that our data set is 2-dimensional with 2 variables *x*, *y* and that the eigenvectors and eigenvalues of the covariance matrix are as follows:

$$\boldsymbol{v_1} = \begin{bmatrix} \mathbf{0.6778736} \\ \mathbf{0.7351785} \end{bmatrix} \lambda_2 = \mathbf{1.284028}$$

$$\boldsymbol{v_2} = \begin{bmatrix} \mathbf{-0.7351785} \\ \mathbf{0.6778736} \end{bmatrix} \lambda_2 = \mathbf{0.04908323}$$

If we rank the eigenvalues in descending order, we get $\boldsymbol{\lambda_1} > \boldsymbol{\lambda_2}$, which means that the eigenvector that corresponds to the first principal component (PC1) is v_1 and the one that corresponds to the second component (PC2) is $\boldsymbol{v_2}$.

After having the principal components, to compute the percentage of variance (information) accounted for by each component, we divide the eigenvalue of each component by the sum of eigenvalues. If we apply this on the example above, we find that PC1 and PC2 carry respectively 96% and 4% of the variance of the data.

✓ **Step 4: Feature Vector**

As we saw in the previous step, computing the eigenvectors and ordering them by their eigenvalues in descending order, allow us to find the principal components in order of significance. In this step, what we do is, to choose whether to keep all these components or discard those of lesser significance (of low eigen values), and form with the remaining ones a matrix of vectors that we call *Feature vector*.
So, the feature vector is simply a matrix that has as columns the eigenvectors of the components that we decide to keep. This makes it the first step towards dimensionality reduction, because if we choose to keep only $\boldsymbol{p}$ eigenvectors (components) out of $\boldsymbol{n}$, the final data set will have only $\boldsymbol{p}$ dimensions. [37]

To show the feature vector process we continue with the example from the previous step, we can either form a feature vector with both of the eigenvectors $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$:

$$\begin{bmatrix} \mathbf{0.6778736} & \mathbf{-0.7351785} \\ \mathbf{0.7351785} & \mathbf{0.6778736} \end{bmatrix}$$

Or discard the eigenvector $\boldsymbol{v_2}$, which is the one of lesser significance, and form a feature vector with $\boldsymbol{v_1}$ only:

$$\begin{bmatrix} \mathbf{0.6778736} \\ \mathbf{0.7351785} \end{bmatrix}$$

Discarding the eigenvector $\boldsymbol{v_2}$ will reduce dimensionality by 1, and will consequently cause a loss of information in the final data set. But given that v_2 was carrying only 4% of the information, the loss will be therefore not important and we will still have 96% of the information that is carried by $\boldsymbol{v_1}$.

✓ **Last Step: Recast the data along the Principal Components Axes**

In the previous steps, apart from standardization, you do not make any changes on the data, you just select the principal components and form the feature vector, but the input data set remains always in terms of the original axes (i.e, in terms of the initial variables).

In this step, which is the last one, the aim is to use the feature vector formed using the eigenvectors of the covariance matrix, to reorient the data from the original axes to the ones represented by the principal components (hence the name Principal Components Analysis). This can be done by multiplying the transpose of the original data set by the transpose of the feature vector. [37]

$$\boldsymbol{Final\ DataSet} = \boldsymbol{Feature\ Vector^T} * \boldsymbol{Standardized\ Original\ Data\ Set^T}$$

C. Applications of Principal Component Analysis:

PCA is predominantly used as a dimensionality reduction technique in domains like facial recognition, computer vision and image compression. It is also used for finding patterns in data of high dimension in the field of finance, data mining, bioinformatics, psychology, etc [47]

III.4.2 Krylov Subspace Methods

A. Definition

Krylov subspace methods have become a very useful and popular tool for solving large sets of linear and non-linear equation, and large eigenvalue problems. One of reason for their popularity is their simplicity and their generality. These methods have been increasingly accepted as efficient and reliable alternative to the more expensive methods that are usually employed for solving dense problems. This trend is likely to accelerate as models are becoming more complex and give rise to larger and larger matrix problems. [38]

B. Why Krylov Methods?

How do you solve a system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ when your coefficient matrix A is large and sparse (i.e., contains many zero entries)?

What if the order n of the matrix is so large that you cannot afford to spend about $\boldsymbol{n^3}$ operations to solve the system by Gaussian elimination? Or what if you do not have direct access to the matrix? Perhaps the matrix $\boldsymbol{A}$ exists only implicitly as a subroutine that, when given a vector $\boldsymbol{v}$, returns $\boldsymbol{Av}$.

In this case you may want to use a Krylov method. Krylov methods are used in numerical as well as in symbolic computation. Since there is no universally agreed upon definition, we say here that a Krylov method solves $\boldsymbol{Ax} = \boldsymbol{b}$ by repeatedly performing matrix-vector multiplications involving $\boldsymbol{A}$ (this excludes methods like Lanczos biorthogonalization, QMR (Quality Management Representative), and biconjugate gradient methods that also require matrix-vector multiplications involving the conjugate transpose $\boldsymbol{A}^*$ starting with an initial guess $\boldsymbol{x_0}$, a Krylov method bootstraps its way up to ever more accurate approximations $\boldsymbol{x_k}$ to a desired solution. In iteration k a Krylov method produces an approximate solution $\boldsymbol{x_k}$ from a Krylov space generated by a vector $\boldsymbol{c}$, [48]

$$\boldsymbol{k_k(A, c) \equiv span\{c, Ac, \dots, k^{k-1}\}} \tag{3.23}$$

Popular choice is $\boldsymbol{c = b}$ (because one can obtain convergence estimates, and because there is often no other problem-dependent guess) and $\boldsymbol{x_0 = 0}$ (we deal with a non zero$\boldsymbol{x_0}$). That's why we restrict ourselves to Krylov spaces $\boldsymbol{k}$, $\boldsymbol{k(A, b)}$ that are generated by the right-hand side of a linear system$\boldsymbol{Ax = b}$.

Let's look at a specific example.

C. An Example of a Krylov Method

Power iteration: find vectors$\boldsymbol{x^{(k)}}$; each is in the subspace spanned by $\boldsymbol{A^k x^{(0)}}$.

Span of $\{\boldsymbol{v_1}, \dots, \boldsymbol{v_m}\} = \{\boldsymbol{c_1 v_1 + c_2 v_2 + c_m v_m : c_m, \dots, c_m} \in ₵\}$

$< \boldsymbol{v_1, v_2, \dots, v_m} >$

$< \boldsymbol{x^{(k)}} >= \{\boldsymbol{cx^{(k)}} : \boldsymbol{c} \in ₵\}$ One dimensional

We have no real choice of the "best possible" eigenvector that lies in a one-dimensional space.

Suppose we looked in the $\boldsymbol{k+1}$ dimensional space spanned by all of the previous iteration:

$< \boldsymbol{x^{(0)}, Ax^{(0)}, \dots, A^{k-1}x^{(0)}, A^k x^{(0)}} >$

$\boldsymbol{A^{k-1}x^{(0)}}$is previeus iterations

The answer can only get better!

(Modify notation) same as the column space of the matrix

$$\boldsymbol{k_m = [u, Au, \dots, A^{m-1}u]}$$

$\boldsymbol{k_m}$ is Krylov matrix $\boldsymbol{n \times m}$ with $\boldsymbol{n \geq m}$

And depends on A and starting vector u.

$$\textbf{Krylov space} =< \boldsymbol{u, Au}, \dots, \mathbf{A^{m-n}u} >$$

Here, $\boldsymbol{A}$ is $\boldsymbol{n \times n}$.

Some facts about these spaces:

- Nested $\boldsymbol{k_1 \subset k_2 \subset k_3 \subset \cdots}$

$$\mathbf{C1u + OAu + OA2u} \in \mathbf{k_3}$$

- Can be constructed by repeated applications of A times vector.
  $\mathbf{u, Au, A(Au), A(A^2u)}$ etc.
- If x is in $\boldsymbol{k_m}$ then $x = k_m z$ for some m-vector z.

$\boldsymbol{x = z_1 u + z_2(Au) + \cdots + z_m(A^{m-1}u) = \ k_m z}$

- If A is nonsingular, $\boldsymbol{dim = m}$. of $\boldsymbol{k_m}$

our big idea for a while: Replace the action of A over all of $\boldsymbol{R^n}$ with its approximate behavior over the lower dimensional $\boldsymbol{k_m}$, for $\boldsymbol{m = 1, 2, 3}$, ...

First question: How?

First answer: consider linear system, $\boldsymbol{A_x = b \quad \rightarrow b - A_x = 0}$

Full problem: $\boldsymbol{\min\limits_{x \in R^n} \|b - A_x\|_2}$

The solution of this problem is: $\boldsymbol{x = A^{-1} b}$

Approximate problem: $\boldsymbol{\min\limits_{x \in k_m} \|b - A_x\|_2}$

$\boldsymbol{k_m}$is smaller space

Equivalent to $\boldsymbol{x = k_m z}$, so

$$\boldsymbol{\min_{z \in R^m} \|b - A(k_m z)\|_2}$$

$\boldsymbol{\min\limits_{z} \|b - A(k_m)z\|_2}$ **:** Ordinary linear least squares problem

III.5 Conclusion

Although PCA in its standard form is a widely used and adaptive descriptive data analysis tool, it also has many adaptations of its own that make it useful to a wide variety of situations and data types in numerous disciplines. Adaptations of PCA have been proposed, among others, for binary data, ordinal data, compositional data, discrete data, symbolic data or data with special structure, such as time series [41] or datasets with common covariance matrices [42]

It is also the techniques described in the previous section, and other recent developments elsewhere, suggest that in many of the problems in control, one can work in a Krylov subspace of reasonably small dimension [38]

Many of the optimization techniques in control can be carried out by replacing the fulln-dimensional variable x by an approximation derived from replacing x by its m dimensional approximation. This may open up interesting new approaches and theoretical questions as to the accuracy of the corresponding approximations [38]

Chapter IV

VI.1 Introduction:

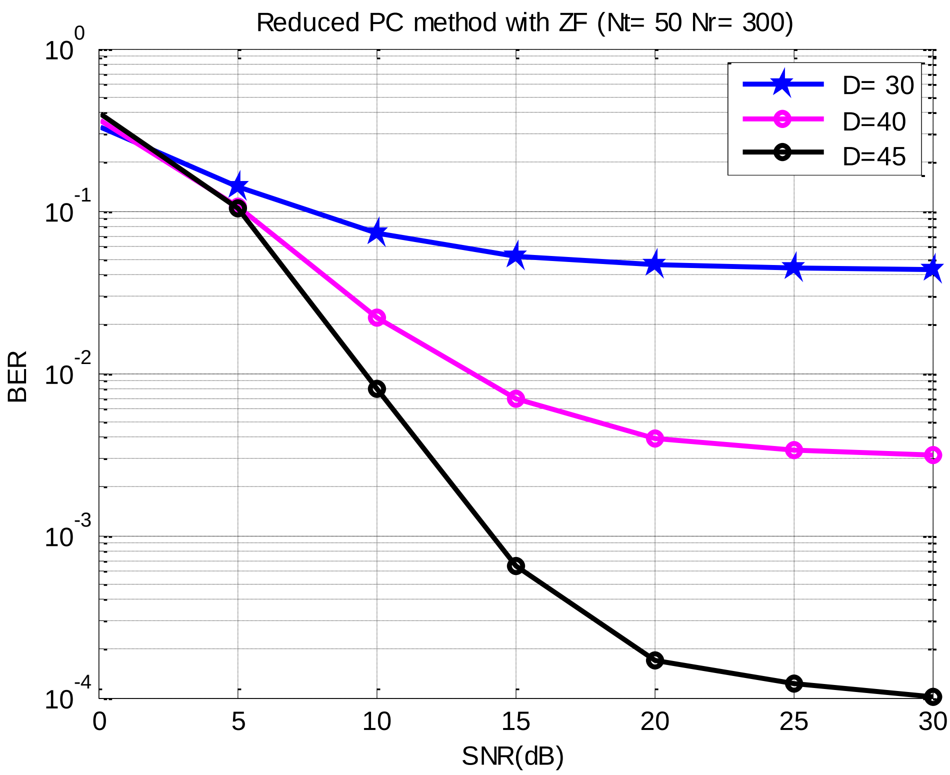
In this chapter we will present the effect of reducing complexity of massive MIMO detectors. Based on the algorithm of Principal Component method (PC) and Krylov Subspace method (KS) as an example, we will illustrate different simulations by varying different parameters such as the parameter of reducing $\boldsymbol{D}$ and the number of transmitted and received antennas $(\boldsymbol{N_t})$ and $(\boldsymbol{N_r})$ and show their impact on the performance of the system by measuring the BER.

In the following simulations different number of transmitted antennas $\boldsymbol{N_t}$and received antenna $\boldsymbol{N_r}$with reducing parameter $\boldsymbol{D}$ are taken**.** A BPSK modulation and a flat fading Rayleigh massive MIMO channel are considered.

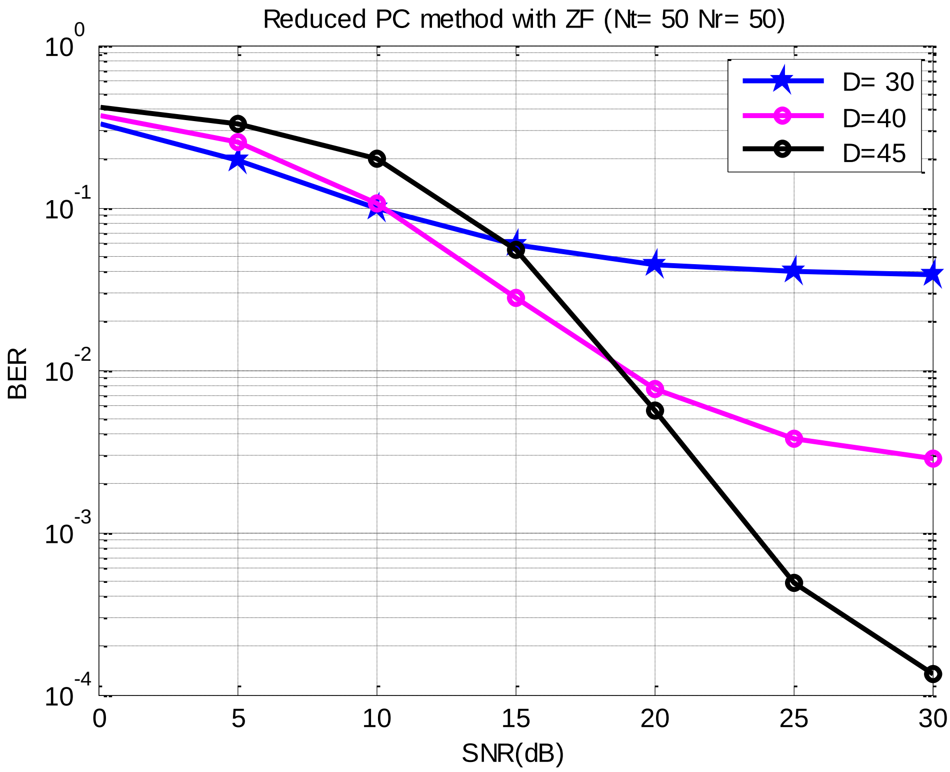
VI.2 Reducing complexity of massive MIMO detector based on PC method:

The following graphs show the performance of massive MIMO system by giving Bit Error Rate (BER) versus Signal to Noise ratio (SNR). Different parameters of $\boldsymbol{D}$ =30,40and 45, where $\boldsymbol{D}$is less than $min(\boldsymbol{N_t}, \boldsymbol{N_r})$are used.

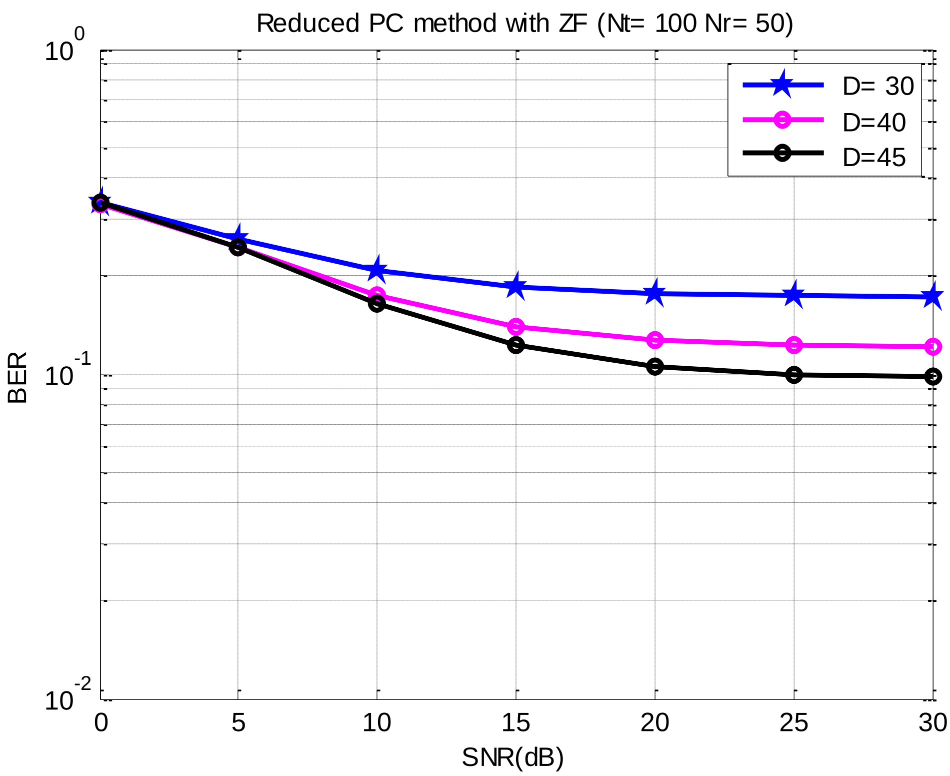
Using PC method is serve to reduced dimension of received vector $\boldsymbol{r}[\boldsymbol{i}]to\boldsymbol{r_D}[\boldsymbol{i}]$**;** with $\boldsymbol{r_D}[\boldsymbol{i}] = \boldsymbol{T^H}_{\boldsymbol{D}\times\boldsymbol{N_r}}\boldsymbol{r}[\boldsymbol{i}]$, where $\boldsymbol{T}$corresponds to a unitary matrix whose columnsare the $\boldsymbol{D}$eigenvectors corresponding to the $\boldsymbol{D}$largest eigenvectors of an estimate of the covariance matrix $\boldsymbol{R}$.



Figure(IV.1): BER versus SNR, with$(\boldsymbol{N_t}, \boldsymbol{N_r})$= (50,300) and different value of $\boldsymbol{D}$



Figure(IV.2): BER versus SNR, with $(\boldsymbol{N_t}, \boldsymbol{N_r})$ = (50, 50) and different value of ***D***



Figure(IV.3): BER versus SNR, with $(\boldsymbol{N_t}, \boldsymbol{N_r})$ = (100,50) and different value of ***D***

The Figures: (IV.1), (IV.2) and (IV.3) illustrate the BER of ZF detector with reduced dimension matrix $\boldsymbol{T}$in function of different value of SNR and received antennasNr. We remark that when $Nr \gg Nt$ the performance is better than when$Nr = Nt$, but the worse case is when$Nr \ll Nt$. These because the high number of received antenna is required to create more diversity and help to reduce the BER.

Greater value of $\boldsymbol{D}$istaken (near to the min (Nt, Nr)), smaller BER is given but with more complexity and the reducing process is quite important in this case. A situation is near to full complexity of ZF detector.

VI.3 Reducing complexity of massive MIMO detector based on Krylov subspace method:

The following graphs show the performance of massive MIMO system by giving Bit Error Rate (BER) versus Signal to Noise ratio (SNR). Different parameters of $\boldsymbol{D}$ =3 and 10, where $\boldsymbol{D}$is less than $min(\boldsymbol{N_t}, \boldsymbol{N_r})$are used.

$\boldsymbol{T}_{\boldsymbol{D}\times\boldsymbol{N_r}}[\boldsymbol{i}] = [\boldsymbol{t_k}[\boldsymbol{i}], \boldsymbol{R}[\boldsymbol{i}]\boldsymbol{t_k}[\boldsymbol{i}], \dots, \boldsymbol{R}^{\boldsymbol{D}-\boldsymbol{1}}[\boldsymbol{i}]\boldsymbol{t_k}[\boldsymbol{i}]]$; $\boldsymbol{t_k}[\boldsymbol{i}] = \boldsymbol{R}[\boldsymbol{i}]$ the covariance $\boldsymbol{k} = \boldsymbol{1}, \boldsymbol{2} \dots \boldsymbol{D}$correspond to the bases of the Krylov subspace.

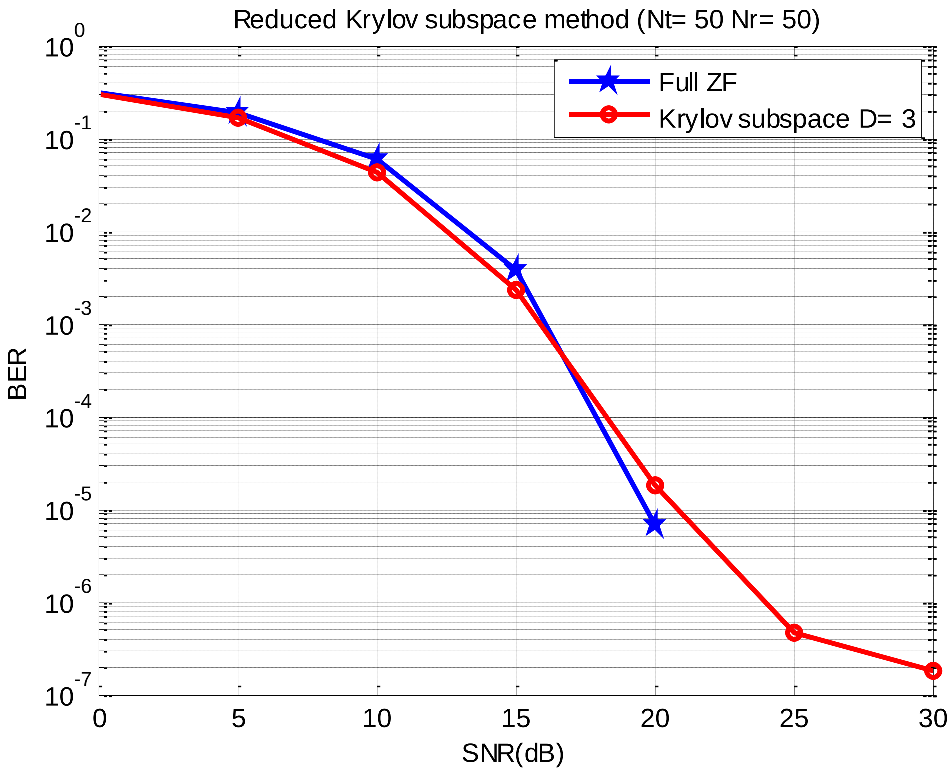


Figure (IV.4): BER versus SNR of ZF detector and Krylov Subspace, with $(\boldsymbol{N_t}, \boldsymbol{N_r})$ = (50, 50) and $\boldsymbol{D} = \boldsymbol{3}$

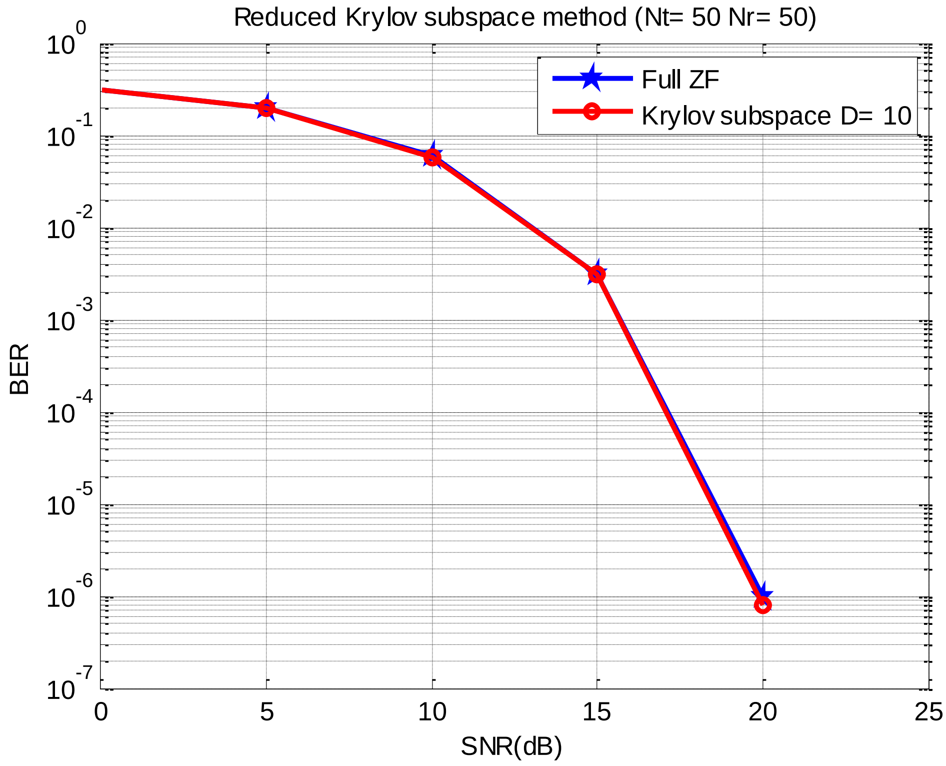


Figure (VI.5): BER versus SNR of ZF detector and Krylov Subspace, with (N_t, N_r) = (50,50) and $D = 10$

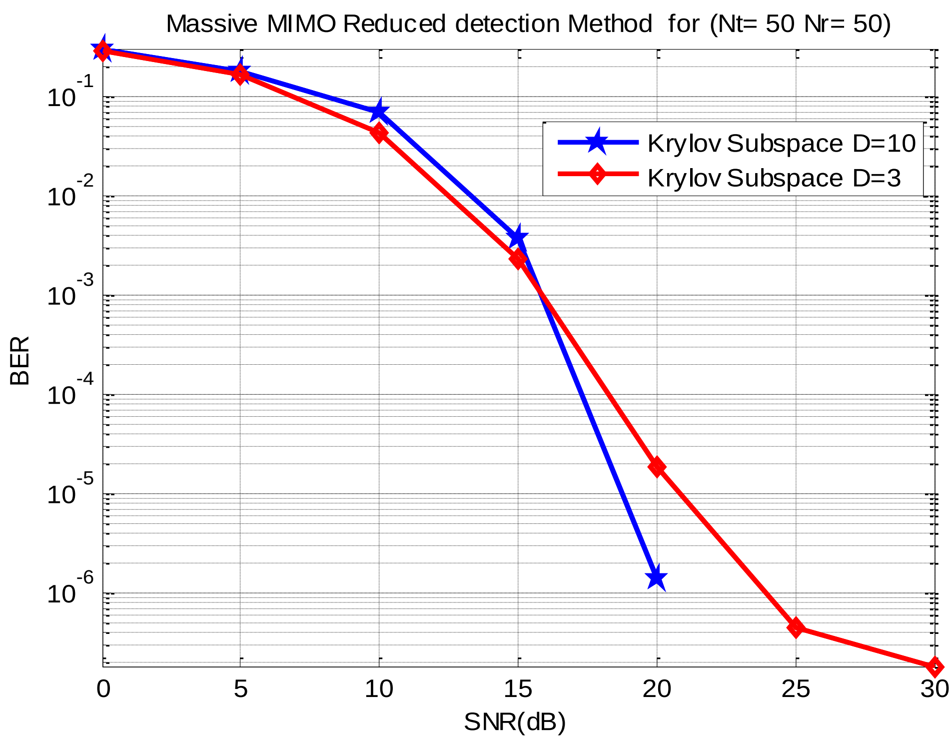


Figure (VI.6): BER versus SNR of Krylov Subspace, with (N_t, N_r) = (50,50) and D =3,10

We observe through (Figures IV.4,5 and 6) that with SNR is less than 5, Krylov subspace method gives the same full ZF performance and gives a lower error rate when SNR is greater. The Krylov subspace gives a same performance as ZF detector full complexity with very high reduction complexity.

VI.4 Conclusion:

In this chapter we present two method of reducing complexity in massive MIMO detection: Principal Component method and Krylov subspace.

We conclude that the Krylov subspace method has more performance than Principal Component but it works only when $N_t = N_r$ with more complexity of calculation

General Conclusion

The need to accommodate more users at higher data rates with better reliability while consuming less energy has forced the birth of a new generation of 5G mobile communication are based, among other things, on the massive MIMO technique, where hundreds of antennas are used. This work is devoted to study some algorithms that can reduce the complexity of detectors as ZF and MMSE who base on inversing the channel matrix of a massive MIMO in 5G system.

In first Chapter, we presented the key pillars building the 5G, and explain its roles and importance for achieving the 5G; the communication systems of the next generation evolve towards greater flexibility in different aspects. Increased flexibility is the key to meeting the various requirements. The 5G enabling technologies and applications will facilitate our life style by connecting everything on internet, and offering faster speeds internet the previous generations.

Then, in chapter 2, MIMO technology is logically classified into one of three categories, whose development occurred during roughly disjoint epochs: Point-to-Point MIMO, Multiuser MIMO, and Massive MIMO.

A particular focus on massive MIMO systems was presented. Review the benefits and challenges of this technology. Massive MIMO marks a clear break with current practice by using a significant excess of service antennas on active terminals. Additional antennas allow energy to be concentrated in smaller and smaller regions to significantly improve throughput and radiated energy efficiency.

Chapter 3 was dedicated for signal detection and parameter estimation techniques for multiuser multiple-antenna wireless systems with a very large number of antennas, known as massive multi-input multi-output (MIMO) systems. The effect of reducing complexity of massive MIMO detectors, based on two methods principal component (PC) and Krylov subspace.

The last chapter is a presentation of different simulations of massive MIMO detectors such as ZF and MMSE, then reduced algorithms as Principal Component (PC) and Krylov subspace method were taken as an example to show how these kind of methods can simplifying the complexity of calculation of inverse matrix channel. The performance of these algorithms is similar to that of full complexity detectors (ZF and MMSE) but in favorite conditions like: High rapport N_r/N_t for PC method and dominant diagonal matrix channel for Krylov subspace.

Bibliography

[1] Moret, Boris. Amplificateur de puissance autonome pour applications OFDM et beamforming de la 5G aux fréquences millimétriques en technologie CMOS avancée. Bordeaux, 2017.

[2] *Enalblingtechnolgies* .(2019).

[3] DEBBAH Nassim,ATMANI Sofiane. « Gestion de la mobilité dans les réseaux cellulaires ». Université ABDERRAHMANE MIRA BEJAIA

[4] DACI kaoutar,DOUwarda , « Etude de Massive MIMO pour 5G » , université Echahid Hamma Lakhdar EL-Oued

[5] Xiang, Wei, Kan Zheng, and Xuemin Sherman Shen, eds. 5G mobile communications. Springer, 2016.

[6] Sons, rodrigez, J. (Éd.). fundamental of 5g mobile network

[7] HAJRI, Salah Eddine. (Thèse de doctorat 2018). L'amélioration des performances des systèmes sans fil 5G par groupements adaptatifs des utilisateurs. Université Paris-Saclay.

[8] LIEN, Shao-Yu, SHIEH, Shin-Lin, HUANG, Yenming, et al. 5G new radio: Waveform, frame structure, multiple access, and initial access. IEEE communications magazine, 2017, vol. 55, no 6, p. 64-71.

[9] 5G Frame Structure. AUGUST 7, 2018[En ligne]. http://spectracells.com/2018/08/ 07/5g-frame-structure/ [consulté le 20 Mai 2019].

[10] 5G NR Terminologies – Subcarrier Spacing, Fram-Subframe, Slot and Symbol. JUNE 10, 2018. [En ligne]. http://www.techplayon.com/understanding-basic-5g-nr-terminologies-subcarrier-spacing-frame-and-subframe-slot-and-ofdm-symbols/ [consulté le 20 Mai 2019].

[11] 5G/NR-FrameStructure[Enligne]. https://www.sharetechnote.com/html/5G/5G_Fr-ameStructure.html [consulté le 20 Mai 2019].

[12] NR Resource Block Definition and RBs Calculation. April 24,2019 [En ligne]. http://www.techplayon.com/nr-resource-block-definition-and-rbs-calculation/[consulté le 27 Mai 2019].

[13] CHALLITA, Frédéric, MARTINEZ-INGLES, Maria-Teresa, LIÉNARD, Martine, et al. Line-of-Sight Massive MIMO Channel Characteristics in an Indoor Scenario at 94 GHz. IEEE Access, 2018, vol. 6, p. 62361-62370.

[14] Study Paper on: Multiple-Input Multiple-Output (MIMO) Technology.

[15] NASRI Amel; BENFETIMA Amel. Study and Simulation of the Transmit Diversity Technique in LTE Systems. 2017.

[16] Fundamentals of Massive MIMO / aut. Erik G.larssonthomas L marzetta, hongyong, Hilenquoc Ngo. - 2016.

[17] Abboud, Ahmad. Interference mitigation in 5G mobile networks: Uplink pilot contamination in TDD massive MIMO scheme. Diss. Limoges, 2017.

[18] Hien Quoc Ngo, ‘’Massive MIMO: fundamentals and system Desings’’, Printed in Sweden by LiU-Tryck, Linköping 2015

[19] Massive MIMO Detection Techniques: A Survey / Mahmoud A. Albreem, Senior Member, IEEE, Markku Juntti, Senior Member, IEEE,and Shahriar Shahabuddin, Member, IEEE

[20] J. Chen, H. Chen, H. Zhang, and F. Zhao, “Spectral-energy efficiencytradeoff in relay-aided massive MIMO cellular networks with pilotcontamination,” IEEE Access, vol. 4, pp.5234–5242, July 2016.

[21] D. Kapetanovic, G. Zheng, and F. Rusek, “Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks,” IEEE Commun. Mag., vol. 53, no. 6, pp. 21–27, June 2015

[22] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “MassiveMIMO for next generation wireless systems,” IEEE Commun. Mag.,vol. 52, no. 2, pp. 186–195, February 2014.

[23] H. Alshamary, “Coherent and non-coherent data detection algorithmsin massive MIMO,” Master’s thesis, University of Iowa, 2017.

[24] E. Bj¨ornson, E. G. Larsson, and T. L. Marzetta, “Massive MIMO: tenmyths and one critical question,” IEEE Commun. Mag., vol. 54, no. 2,pp. 114–123, February 2016.

[25] S. Jin, X. Wang, Z. Li, K. K. Wong, Y. Huang, and X. Tang, “Onmassive MIMO zero-forcing transceiver using time-shifted pilots,”IEEE Trans. Veh. Technol., vol. 65, no. 1, pp. 59–74, Jan 2016.

[26] E. Bj¨ornson, J. Hoydis, and L. Sanguinetti, “Massive MIMO has capacity,” IEEE Trans. Wireless Commun., vol. 17, no. 1,pp. 574–590, Jan 2018.

[27] Y. G. Lim, C. B. Chae, and G. Caire, “Performance analysis of massiveMIMO for cell-boundary users,” IEEE Trans. Wireless Commun.,vol. 14, no. 12, pp. 6827–6842, Dec 2015.

[28] A. A. Lu, X. Gao, Y. R. Zheng, and C. Xiao, “Low complexity polynomial expansion detector with deterministic equivalents of the moments of channel Gram matrix for massive MIMO uplink,” IEEE Trans. Commun., vol. 64, no. 2, pp. 586–600, Feb 2016

[29] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer, “Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations,” IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 916–929, Oct 2014.

[30] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu, and F. Rusek, “Hardware efficient approximative matrix inversion for linear pre-coding in massive MIMO,” in Proc. IEEE Int. Symp. on Circuits and Systems, June 2014, pp. 1700–1703.

[31] H. Prabhu, J. Rodrigues, O. Edfors, and F. Rusek, “Approximative matrix inverse computations for very-large MIMO and applications to linear pre-coding systems,” in Proc. IEEE Wireless Commun. And Networking Conf., April 2013, pp. 2710–2715.

[32] I. Al-Nahhal, M. Alghoniemy, A. B. A. El-Rahman, and Z. Kawasaki, “Modified zero forcing decoder for ill-conditioned channels,” in 2013 IFIP Wireless Days (WD), Nov 2013, pp. 1–3.

[33] I. Al-Nahhal, M. Alghoniemy, O. Muta, and A. B. A. El-Rahman, “Reduced complexity k-best sphere decoding algorithms for ill-conditioned mimo channels,” in IEEE Annual Cons. Commun. Netw. Conf., Jan 2016, pp. 183–187.
[34] A Robust Hybrid Iterative Linear Detector for Massive MIMO Uplink Systems
Mahmoud A. Albreem , Mohammed H. Alsharif and Sunghwan Kim , 23 January 2020; Accepted: 17 February 2020; Published: 21 February 2020.

[35] Albreem, M. Efficient initialisation of iterative linear massive MIMO detectors using a stair matrix. Electron. Lett. 2020, 56, 50–52.

[36] Jiang, F.; Li, C.; Gong, Z.; Su, R. Stair Matrix and its Applications to Massive MIMO Uplink Data Detection. IEEE Trans. Commun. 2018.

[37] Zakaria jaadi,A STEP BY STEP EXPLANATION OF PRINCIPAL COMPONENT ANALYSIS, September 4, 2019 Updated: May 19, 2020

[38] Youcef Saad, ‘’Overview of Krylov Subspace methods with applications to control problems’’

[39] B. Shim and I. Kang, “On further reduction of complexity in tree pruning based sphere search,” IEEE Trans. Commun., vol. 58, no. 2, pp. 417–422, Feb. 2010.

[40] Rodrigo C. de Lamare and Raimundo Sampaio-Neto, "Signal Detection and Parameter Estimation in Massive MIMO Systems", Centre for Telecommunications Studies (CETUC), Pontifical Catholic University of Rio de Janeiro, Gavea - 22453-900, Rio de Janeiro, Brazil, Communications Research Group

[41] Jolliffe IT. 2002. *Principal component analysis*, 2nd edn New York, NY: Springer-Verlag

[42] Hallin M, Paindaveine D, Verdebout T. 2014. Efficient R-estimation of principal and common principal components. J. Am. Stat. Assoc. 109, 1071–1083. (10.1080/01621459.2014.880057)

[43] R.C. de Lamare and R. Sampaio-Neto, “Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems,” IEEE Trans. Veh. Technol., vol. 60, no. 6, pp.2482-2494, July 2011.

[44] Raimundo_Sampaio-Neto. Rodrigo De Lamare,‘’Detection and Estimation Algorithms in Massive MIMO Systems’’ · August 2014 *with* 976 Reads

[45] S. Haykin, Adaptive Filter Theory, 4th ed. Englewood Cliffs, NJ: Prentice- Hall, 2002.

[46] R. C. de Lamare, P. S. R. Diniz, “Set-Membership Adaptive Algorithms Based on Time-Varying Error Bounds for CDMA Interference Sup-pression”, IEEE Transactions on Vehicular Technology, vol.58, no.2, pp.644-654, Feb. 2009.
[47] https://www.dezyre.com/data-science-in-python-tutorial/principal-component-analysis-tutorial.
[48] The Idea Behind Krylov Methods. Ilse C.F. Ipsen and Carl D. Meyer
[49] Performance Analysis and Location Optimization for Massive MIMO Systems with Circularly Distributed Antennas.