Universidade de Aveiro
2017

Departamento de Eletrónica, Telecomunicações e Informática

André Daniel Carvalho Lopes

Channel Quantization Techniques for Hybrid Massive MIMO Millimeter Wave Systems

Técnicas de Quantização para Sistemas de Comunicação Híbridos na Banda de ondas Milimétricas com um número elevado de antenas

Universidade de Aveiro
2017

Departamento de Eletrónica, Telecomunicações e Informática

André Daniel Carvalho Lopes

Channel Quantization Techniques for Hybrid Massive MIMO Millimeter Wave Systems

Técnicas de Quantização para Sistemas de Comunicação Híbridos na Banda de ondas Milimétricas com um número elevado de antenas

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Electrónica e Telecomunicações, realizada sob a orientação científica do Professor Doutor Adão Silva (orientador), Professor Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro e da Doutora Sara Teodoro (co-orientadora), investigadora do Instituto de Telecomunicações de Aveiro

o júri / the jury

presidente / president

Prof. Doutor Aníbal Manuel de Oliveira Duarte
Professor Catedrático, Universidade de Aveiro

vogais / examiners committee

Prof. Doutor Fernando José da Silva Velez
Professor Auxiliar, Dep. De Engª Electromecânica da Fac. De Engenharia da Univ. da Beira Interior

Prof. Doutor Adão Paulo Soares da Silva
Professor Auxiliar, Universidade de Aveiro

agradecimentos

Queria deixar um agradecimento a todas as pessoas que, de uma forma ou de outra, estiveram presentes neste trajeto.

Destaco de forma significativa:

Os meus familiares: pais, tios, avós e irmão por terem depositado em mim total apoio e confiança.

Os professores Adão Silva e Sara Teodoro, por me terem acompanhado e orientado de forma criteriosa ao longo deste último ano.

Por último, mas não menos importante, a todos os meus amigos que estiveram sempre presentes durante esta jornada académica.

A todos, um Muito Obrigado.

palavras-chave

Arquitectura hibrida , 5G, MIMO massivo , ondas milimétricas, precodificação/equalização para multi-antenas, quantização de canal, quantização uniforme, sistemas OFDM.

resumo

Desde o aparecimento das comunicações móveis, os utilizadores desta tecnologia têm vindo a crescer exponencialmente todos os dias. A escalada do crescimento do tráfego móvel foi imposta, principalmente, pela proliferação de *smartphones* e *tablets*. O uso crescente e intensivo das comunicações sem fios pode levar no futuro a um ponto de rutura, onde os sistemas tradicionais não suportam a capacidade requerida, a eficiência espectral e eficiência enérgica. Por outro lado, para cobrir toda esta necessidade atual de ter mais e mais dados, é necessário fornecer taxas de transmissão mais elevadas, em torno dos gigabits por segundo.

Hoje, quase todos os sistemas de comunicações móveis usam espectro na faixa de 300 MHz - 3GHz. É necessário começar a procurar a gama de espectro 3GHz - 300 GHz para aplicações de banda larga móvel. Aqui vamos apresentar as ondas milimétricas, sendo esta uma maneira de aliviar espectro em frequências mais baixas.

Os sistemas baseados em MIMO foram alvo de pesquisa nos últimos 20 anos e agora fazem parte dos padrões atuais. No entanto, para obter mais ganhos, uma visão mais ampla do conceito MIMO prevê o uso de uma grande quantidade de antenas em cada estação base, um conceito referido como massive MIMO.

A combinação simbiótica destas tecnologias levará ao desenvolvimento de um novo sistema de geração denominado 5G.

O desenvolvimento de técnicas de conhecimento da informação do canal no transmissor é muito importante em sistemas *massive MIMO millimeter wave* reais. Nesta dissertação é proposta e avaliada uma estratégia de envio de informação de canal para o transmissor para sistemas *massive MIMO* OFDM híbrido, onde apenas uma parte dos parâmetros associados ao canal são quantificados e transmitidos para o transmissor. A estratégia de *feedback* proposta é baseada numa quantização uniforme das amplitudes de canal, ângulos de partida e de chegada, no domínio do tempo. Depois de serem enviadas, essas informações são usadas para reconstruir o canal geral no domínio da frequência e a matriz da antena de transmissão, que são então usadas para obter os *precoders* híbridos analógico-digitais. Os resultados numéricos mostram que a estratégia de quantificação proposta atinge um desempenho próximo ao obtido caso se conhecesse o canal perfeito no transmissor, com um baixo *overhead* e complexidade

keywords

Hybrid architecture, 5G, massive MIMO, millimeter waves, precoding / equalization for multi-antennas, channel quantization, uniform quantization, OFDM systems.

abstract

Since the appearance of mobile communications, the users of this technology have been growing exponentially every day. The escalating mobile traffic growth it has been imposed by the proliferation of smartphones and tablets. The increasing and more intensive use of wireless communications may lead to a future breaking point, where the traditional systems will fail to support the required capability, spectral and energy efficiency. On the other hand, to cover all this current need to have more and more data it is necessary to provide a new range of data rates around the gigabits per second.

Today, almost all mobile communications systems use spectrum in the range of 300MHz – 3GHz. It is needed to start looking to the range of 3GHz – 300GHz spectrum for mobile broadband applications. Millimeter waves are one way to alleviate the spectrum gridlock at lower frequencies.

MIMO based systems has been researched for the last 20 years and are now part of the current standards. However, to achieve more gains, a grander view of the MIMO concept envisions the use of a large scale of antennas at each base stations, a concept referred as massive MIMO. The symbiotic combination of these technologies and other ones will lead to the development of a new generation system known as the 5G.

The knowledge of the channel state information at the transmitter is very important in real massive MIMO millimeter wave systems. In this dissertation a limited feedback strategy for a hybrid massive MIMO OFDM system is proposed, where only a part of the parameters associated to the link channel are quantized and fed back. The limited feedback strategy employs a uniform-based quantization for channel amplitudes, angle of departure and angle of arrival in time domain. After being fed back, this information is used to reconstruct the overall channel in frequency domain and the transmit antenna array, which are then used to compute the hybrid analog-digital precoders. Numerical results show that the proposed quantization strategy achieve a performance close to the one obtained with perfect full channel, with a low overhead and complexity.

Contents

List of Acronyms

1G	First Generation
2G	Second Generation
3D	Three Dimensional
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AA	Antenna Array
AMPS	Advanced Mobile Phone System
AoA	Angles of Arrival
AoD	Angles of Departure
AWGN	Additive White Gaussian Noise
BER	Bit-error rate
BS	Base Station
CBSM	Correlation Based Stochastic Model
CDI	Channel Direction Information
CDMA	Code Division Multiple Access
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CP	Cyclic Prefix
CSI	Channel State Information
CSIT	Channel State Information at Transmitter
D-Blast	Diagonal- Bell Labs Space-Time Architecture
DFT	Discrete Fourier Transform
DoF	Degrees of Freedom

EDGE	Enhanced Data Rate for Global Evolution
EGC	Equal Gain Combining
EHF	Extremely High Frequency
EMO	European Mobile Observatory
EV-DO	Evolution Data Optimized
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FFC	Federal Communication Commission
FFT	Fast Fourier Transform
FSF	Frequency Selective Fading
GPRS	General Packet Random Service
GSM	Global System for Mobile Communication
HetNet	Heterogeneous Network
HomoNet	Homogeneous Network
HSPA	High Speed Packet Access
ICI	Inter Carrier Interference
ICIC	Inter-Cell Interference Coordination
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
IMT-2000	International Mobile Telecommunication 2000
IP	Internet Protocol
ISI	Inter Symbol Interference
ITU	International Telecommunication Union
LMDS	Local Multipoint Distribution Service
LTE	Long Term Evolution
M2M	Machine to Machine
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output

MMB	Millimeter Mobile Broadband
MMB BS	Millimeter Mobile Broadband Base Station
MMS	Multimedia Messaging Service
MMSE	Minimum Mean Square Error
mmWave	Millimeter Wave
MRC	Maximal Ratio Combining
MS	Mobile Station
MU-MIMO	Multi User MIMO
NLOS	Non-line of Sight
NMT	Nordic Mobile Telephones
NTT	Nippon Telephone and Telegraph
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PMI	Precoder Matrix Indicator
QoS	Quality of Service
RAT	Radio Access Technologies
RF	Radio Frequency
RVQ	Random Vector Quantization
SC	Selection Combining
SC-FDMA	Single Carrier Frequency Division Multiple Access
SFBC	Space-Frequency Block Codes
SHF	Super High Frequency
SIC	Successive Interference Cancellation
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
SMS	Short Message Service
SNR	Signal-Noise Ratio
SQNR	Signal-to-Quantization-Noise Ratio

STBC	Space-Time Block Codes
SU-MIMO	Single User MIMO
SVD	Singular Value Decomposition
SWC	Switched Combining
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
ULA	Uniform Linear Array
UPA	Uniform Planar Array
UMTS	Universal Mobile Telecommunication System
UQ	Uniform Quantization
V-Blast	Vertical -Bell Labs Space-Time Architecture
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network
ZF	Zero Forcing

Chapter 1

Introduction

In this chapter, the evolution of the wireless communications is addressed, since the primordial technologies that revolutionized the world of mobile communications to nowadays. Then, the motivation and objectives are addressed and after that, the future deployments and the main contributions are addressed. Finally, a summary and a discussion of the following chapters are made.

1.1 Evolution of Telecommunications Systems

Figure 1.1 - The terminal development in the past few years [1]

Nowadays it is hard to believe in a world without the facilities provided by wireless communication technologies. Wireless telecommunications are, in fact, the technological basis of our modern societies. Wireless means communication through radio. In 1865, based on the James Clerk Maxwell 1857's manifest about the electromagnetic radiation, Marconi created the radio transmission. Later, in 1901, Marconi made the first communication through Atlantic Ocean. However, it was only half century later, in the late 1940's in the USA and in the1950's in Europe, that the first phone was ever presented. These 'portable' phones were portably maladjusted and awkwardly extravagant [2].

Later, in 1979, in Tokyo, Japan, the first cellular system in the world became operational by Nippon Telephone and Telegraph (NTT). Two years later, this technology reached Europe, more specifically to the countries of northern Europe like Finland, Sweden and Denmark. The system became known as Nordic Mobile Telephones (NMT). In the USA, in 1982, the Advanced Mobile Phone System (AMPS) was deployed [3]. The fact that the NMT was an international system - the device keeps connected to a network when travelling outside the geographical coverage area of the home network (roaming) - gave to the NMT and to mobile phones larger markets, attracting more companies and more investment to the telecommunication business [1]. So, around the 80's, cellular systems were deployed. Known as 1G – the G stands for Generation - from there, there was a big "boom" in the mobile communication systems. The 1G started as a voice only service and the main characteristic of these systems was the use of the analog techniques for the transmission. They were used in large cells and were not efficient in the use of the spectrum. This system evolved increasingly to integrate other services and because of that, a multidirectional evolution happened. Fig.1.2 shows the mobile cellular network evolution through time.

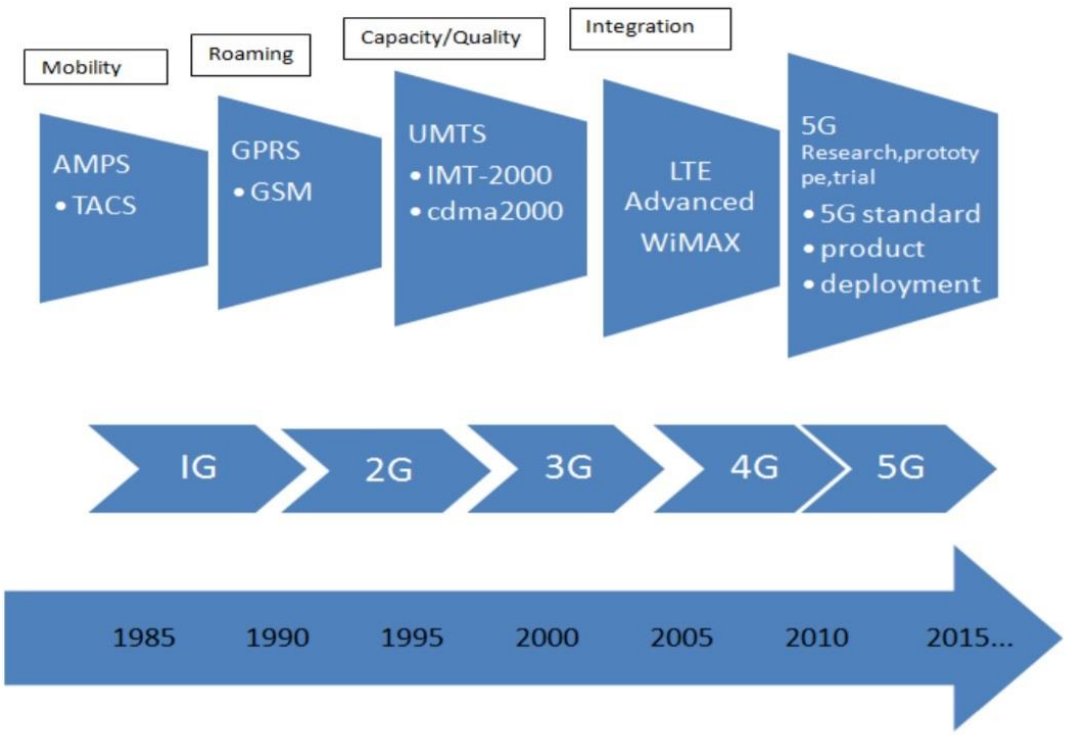

Figure 1.2 - Mobile Cellular Network Evolution Timeline [4]

Back in the 90's, a new Generation was deployed. This second Generation, 2G, like 1G, was one network for one service (voice). However, the transmission technique used changed from analogic to digital. This system use digital multiple access technology

like Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) [3]. TDMA takes into consideration the division of the sign in separated time spaces while CDMA provides to each client an uncommon code to transmit in a multiplex physical channel. Advances in the TDMA technology led to a pan-European standard called Global System for Mobile Communication (GSM). GSM is the most famous and widely used 2G mobile standard. Despite that innovation begins in Europe, this standard is utilized in more than 212 nations on the planet [2]. With the 2G became the instant messages technology called Short Message Service (SMS), picture messages and multimedia messages called Multimedia Messaging Service (MMS) and, compared to first generation, higher spectrum efficiency, better data services and more advanced roaming. The 2G is a success and by far the dominant system.

The continuous search for better performances by the GSM technology led in the mid 90's to the so called 2.5G. In this systems data packet switched domain were introduced. In this evolution was implemented the General Packet Radio Service (GPRS) but, the requirement for higher data rates led to Enhanced Data Rate for Global Evolution (EDGE) [5]. In the EDGE was possible a high volume of the transfer data, but the air-interface that behaves like a circuit switches call make it low efficiently [3].

The letter G became famous in the telecommunication world. The word Generation generally refers to a change in the fundamentals nature of the service. However, the word went only viral in the mid 90's when the studies of 3G started to be known. Fig. 1.3 shows the global distribution of mobile subscribers from 2008 to 2015 by different technology.

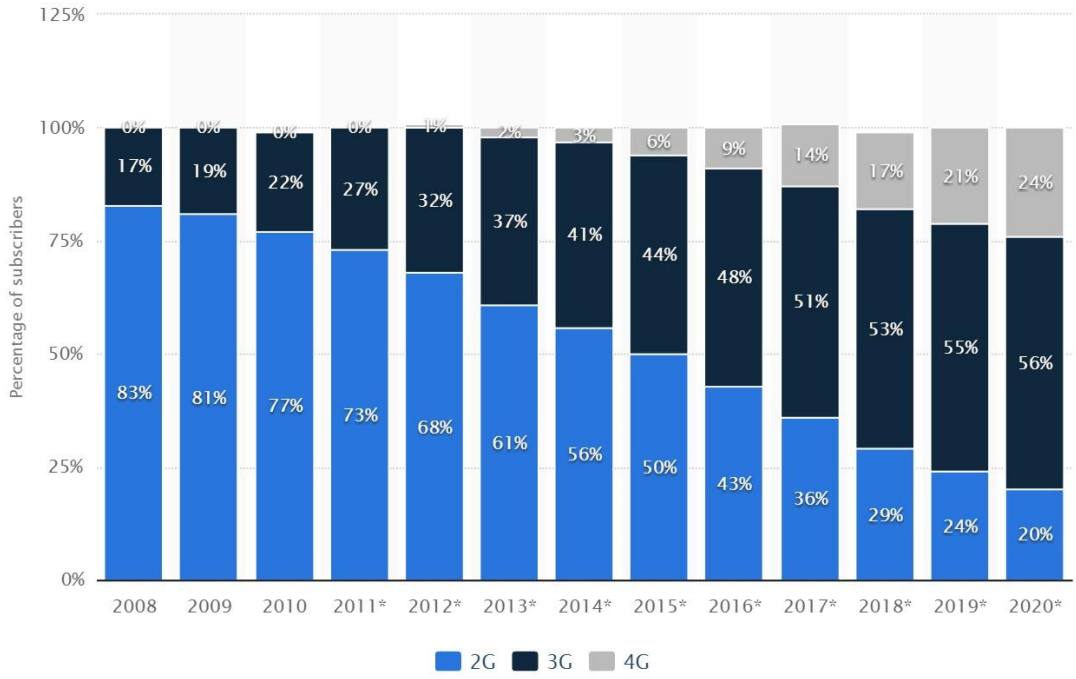

Figure 1.3 - Global distribution of Mobile Subscribers from 2008 to 2015, by technology [6].

The innovation and the investigation on the telecommunication business continued. The pursuit of higher rates and efficiency led to 3G that was introduced in 2000. This new generation's goal was to provide better data rates from 144 Kbps to 384 Kbps in wide coverage areas and up to 2Mbps in local coverage areas [5]. Beside voice communication, 3G includes data services, access to TV and videos, web browsing, video conferencing, navigational maps, etc. However, like in the previous generations, the standard for developing the networks were different around the globe. So, this mobile system was created by a group of companies denominated 3rd Generation Partnership Project (3GPP) using the International Mobile Telecommunications-2000 (IMT-2000) standards which was defined by International Telecommunication Union (ITU). Companies from Europe, North America and Asia were part of the group. 3G is not one standard, but family of standards that can work all together [3]. In Europe was deployed the Universal Mobile Telecommunication System (UMTS) and American variant is called CDMA2000. Wideband CDMA (WCDMA) is the air-interface used in these systems. The first commercialized 3G network was launched in 2001 in Japan. Later, UMTS and CDMA2000 evolved respectively to High Speed Packet Access (HSPA) and Evolution Data Optimized (EV-DO). This evolution became known as 3.5G. The limit of these systems is a real thing. As utilization keeps on an exponential growing, the existing infrastructures are coming to a point of confinement. Cells can be made smaller, allowing reassurance reuse to a point. The following step is an innovation with new transfer speed [2].

For many years, voice calls dominated the traffic in mobile telecommunications networks. But, the years leading to 2010, the escalating mobile traffic growth it's been imposed by the proliferation of smartphones and tablets. The appearance of Apple IPhone and Google Android were an important factor. Related to this, there has also been an evolution in the environment in which mobile systems are deployed and operated in terms of competitions between mobile operators, challenges from other mobile technologies, new regulations in the spectrum use and the market aspects of mobile systems [7]. Around 2010, a new system, the 4G, or the so-called Long Term Evolution (LTE), developed by 3GPP, was presented. The migration from 3G started by when a not so successful first approach technology called WiMAX was developed. This new generation is supported on all Internet Protocol (IP) based systems. The natural next step was those all internet based services also move to the mobile devices, creating what is today known as mobile broadband. A few services were already evolved in the 2.5G systems, but it was not until the 3G and 4G systems that they were deployed [8]. One more reason for the transition to all IP basis is to have a common platform for all the technologies developed so far. The main goal of 4G technology is to provide high speed, high quality, high capacity, security and low cost services for voice and data services, multimedia and internet. It's a technology that tries to fulfill the ubiquity of communication. To achieve this ubiquity, i.e., the capacity to use wireless services anytime and anywhere, terminal mobility is a key factor in 4G. This mobility implies automatic roaming between different wireless networks. So, the 4G integrate different

existing and future wireless technologies to provide freedom of movement and uninterrupted roaming from one technology to another [5]. In LTE is used an Orthogonal Frequency-Division Multiplexing (OFDM) modulation system to downlink and a Single Carrier – Frequency Division Multiple Access (SC-FDMA) in uplink. These multiple access schemes provide orthogonality between users, thus reducing the inter-symbol interference. Another key technology that enables 4G goals above said is the Multiple Input – Multiple Output (MIMO) channel transmission techniques. The coexistence of the LTE and the HSPA provide data rates in mobile systems exceeding the 100 Mbps. But, the demand of more data rates and to meet new requirements from the ITU, in 2008, LTE-Advanced appeared, which theoretically be capable to exceed the 1 Gbps. However, as the demand of higher and higher data rates on the mobile market is always rising, mobile wireless companies must be prepared to support a thousand-fold increase in total mobile traffic until 2020, requiring researchers to seek higher capacities and to find new wireless spectrum beyond the 4G standard [9].

Inside the wireless telecommunication business, LTE has by far the greater support amongst network operators and equipment manufacturers and is likely to be the world's dominant mobile communication technology for some years to come [8].

The rapid evolution of the technology used in telecommunication systems, consumer electronics, and specifically mobile devices has been remarkable in the last 20 years. Moore's law illustrates this and indicates a continuing evolution of processor performance and increased memory size, often combined with reduced size, power consumption, and cost for devices. Combined with a high-speed internet backbone often based on optical fiber networks, we see that a range of technology enablers are in place to go hand-in-hand with advancement in mobile communications technology [7]. These evolutions in the mobile and cellular systems completely changed the paradigm of communication between people.

1.2 An overview towards 5G

Radio technologies have evidenced a rapid and a multidirectional evolution with the launch of the analog cellular systems around 1980. Thereafter, digital wireless communication systems are constantly on a mission to fulfill the growing need of human beings.

If we look back, we will find that every next decade one generation is advancing in the field of mobile technology. Starting from the First Generation, 1G, in the 80's, Second Generation, 2G, in the 90's, Third Generation, 3G, in the 00's, Fourth Generation, 4G, in 10's, and now Fifth Generation, 5G, we are advancing towards to more and more sophisticated and smarter technology.

Even though in 4G there were already high data rates even with a good mobility, the exponential growth of wireless data services caused by mobile internet and smart devices triggered the development of the 5G cellular systems. The European Mobile Observatory (EMO) pointed out that there has been a 92 percent growth in mobile broadband per year since 2006 [10]. As more and more devices go wireless, more and more research need to be addressed.

One of the most important challenges is the physical scarcity of the radio frequencies spectra used for cellular communications. Cellular frequencies use ultra-high-frequencies bands (300 MHz – 3 GHz) for cellular phones. This band is generally referred as *sweet spot* due to its favorable propagation characteristics for commercial wireless applications. These frequencies spectra have been massively used, making difficult for operators to acquire more. It's urgent to look up higher frequencies: Millimeter-Waves (mmWave). Up to now, the portion of the radio frequency (RF) spectrum above 3 GHz, however, has been largely unexploited for commercial wireless applications [11].

Another key aspect of 5G is the development of a grander version of MIMO transmission technique systems. Base stations (BS) equipped with a very large number of antennas (100 or more) that can simultaneously accommodate many co-channel users, an idea referred to as massive MIMO or large antenna arrays.

Individually, each one of these technologies could offer an increase in wireless capacity. However, the combination with one another can envision achieving in capacity the approximate thousand-fold increase in capacity that will be needed in the coming decades [12].

Others specific technical directions have been identified to achieve the optimal system configuration that will maximize success deployment and realization of a powerful wireless world. They are seven and are illustrated in Fig. 1.4. Some are already deployed but the 5th, 6th and 7th are yet under development.

As can be seen in the Fig. 1.4, evolution of the Radio Access Technologies (RATs) was the first improving to be deployed mainly to upgrade the service provision and the cost efficiency. This evolution led us from the Frequency Division Multiple Access (FDMA), Time Division Multiple Access (TDMA) and Wideband Code Division Multiple Access (WCDMA) to Orthogonal Frequency Division Multiple Access (OFDMA).

Along with the past years, the decreasing of the cells size has been a key to improving the capacity and the cost of resources already deployed. There are cells with sizes that range from macro to the most modern small cells.

The 3rd direction illustrated was aimed to the exploitation of a heterogeneous wireless access infrastructures, in terms of RATs. A significant portion of this work has been on the interworking of cellular systems with wireless local area networks (WLANs) [13].

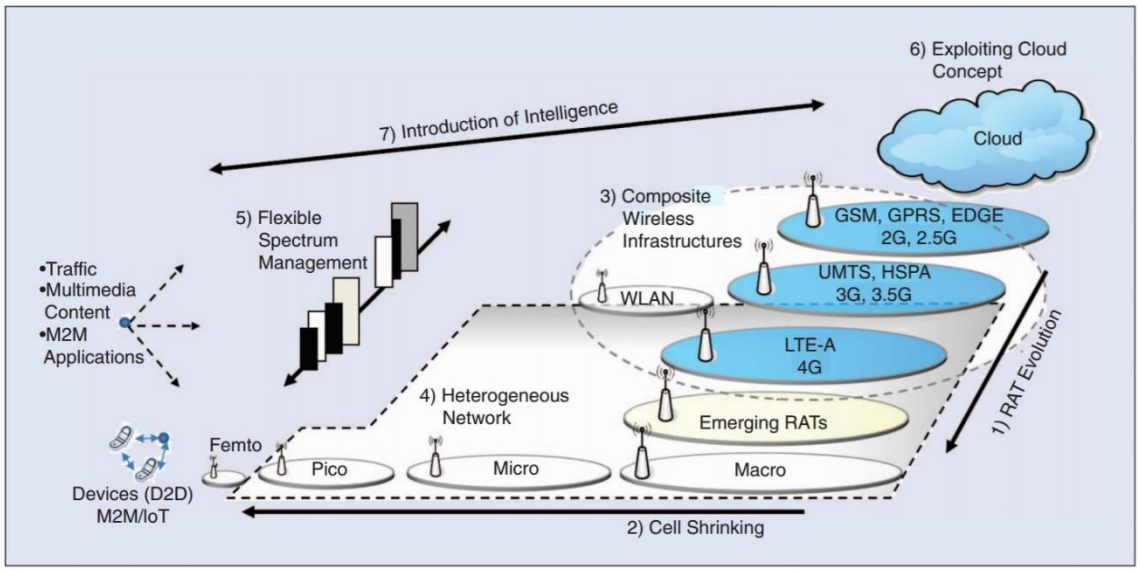

Figure 1.4 - A view of the wireless world; Seven technical directions aiming at the proper applications provisioning, cost-efficient resource provisioning, and the augmentation of the wireless world's intelligence [13]

The 4th direction is related with the heterogeneous networks, which is primarily aimed to increase cost efficiency. These networks may consist on different types of infrastructures elements, such as macro, micro, pico and femto cells. Heterogeneity will also be a feature that is expected to characterize the emerging wireless world, as mixed usage of cells of diverse sizes and access point with different characteristics and technologies in an operating environment are necessary [13].

The 5th way toward 5G is the flexible spectrum management. To improve resource utilization, network operators have the freedom to allocate spectrum to the technologies that they operate and to operate a technology at various spectrum bandslike an opportunistic spectrum access [13].

Another direction, like the Device-to-Device and Machine to Machine Communications are expected to characterize beyond 4G and 5G networks and they are related with the creation of dynamic networking consisting of interconnected end-user equipment or several machines, in the context of the internet of things.

An improvement to capitalize the cost efficiency is the exploitation of the clouds. This reason is that there can be a total cost of ownership savings if wireless networks are based on cloud principles. This is possible through the shared use of storage or computing resources. An achievement in the energy efficiency is also a target.

The ultimate goal is the introduction of the intelligence in the upcoming heterogeneous wireless networks. This concept is essential to specify and provide an efficient handling solutions in the heterogeneous networks and clouds.

5G is the forthcoming revolution of mobile technology and, around 2020, the new mobile network is expected to be deployed. As illustrate in Fig. 1.5, compared with the 4G, it is widely agreed that 5G should achieve 1000 times the system capacity, 10 times the spectra efficiency, energy efficiency and data rates and 25 times the average cell throughput. The aim is to connect the entire world and achieve the perfect and ubiquitous communications between anybody, everybody, wherever they are, whenever they need by whatever electronic device/service/network they wish [14].

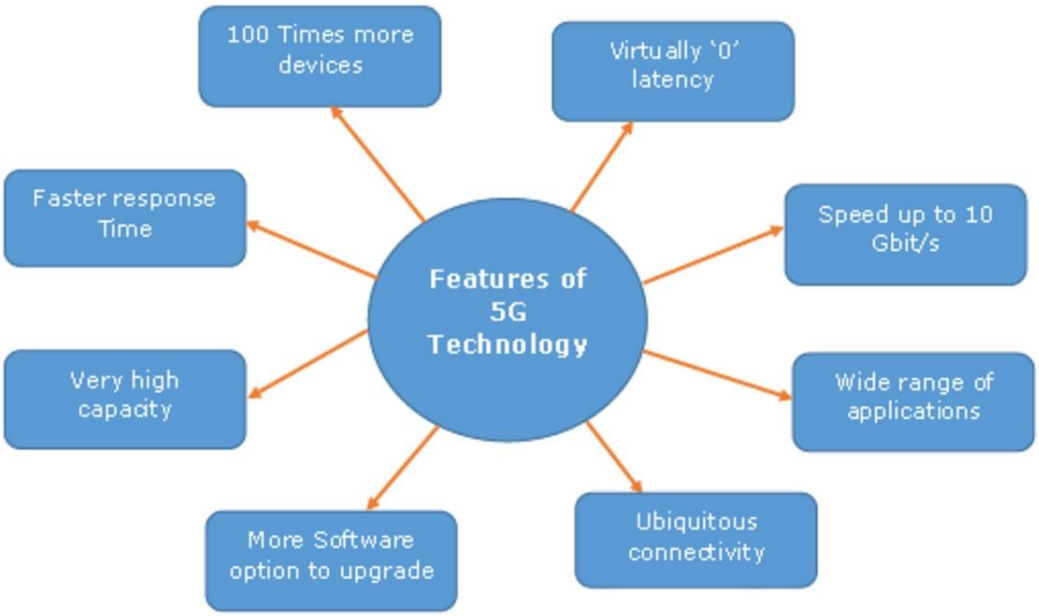

Figure 1.5 - Ideal features of 5G Systems [15]

1.3 Motivation and Objectives

Year after year, the number of mobile communication users increases. In fact, it is known that there are more mobile terminals connected to mobile networks than users, and the number does not stop growing. In addition to this numbers, the requirements of each user increases even more, thus, becoming the pursuit of better data rates, greater capacities, better efficiencies, which are some of the goals of the telecommunications companies that control the market. Until the end of 2016, two thirds of the world's population had a mobile subscription, making a total of almost 5 billion unique subscribers [16].

However, to develop and achieve the goals outlined it is necessary to overcome some barriers. One of them is the overcrowding of the current spectrum used for wireless communications. The 300 MHz – 3 GHz band, known as sweet pot, is completely

outdated and overcrowded and so it is needed to start looking to a new range of frequencies: 3 GHz – 300 GHz. The access to these new frequencies results in the millimeter waves and, from them, it is possible to increase in a grander way the transmission rates.

As mentioned, another concept used to meet the demands of the new generation to come is the massive MIMO. The fact that it is possible to add hundred antennas, or even thousands of them, in one base station will help to improve gains, bandwidth and energy efficiencies, providing the enough received signal power.

Coupled with the massive MIMO concept, the mmWave – due to its size – will allow to allocate in a small space more and more antennas and, consequently, overcome the adjacent difficulties of mmWave propagation. Recent studies show that the combination a large number of antennas in a BS and mobile terminals can offer coverage comparable to the conventional low frequencies systems, but with higher rates [17].

Another advantage of the combination of these two new technologies, mmWave with massive MIMO, is that can exploit new processing technologies different from those used in traditional cellular systems, both on the transmitter and on the receiver side. MIMO precoding/combining in mmWave is generally different [18]. Full digital beamforming requires an RF chain per antenna. This is very unpractical in massive MIMO technologies since there is a huge number of antennas, thus, the complexity of the process will increase. So, to attenuate the full digital complexity, the solution is using a number of RF chains lower than the number of antennas and thus designing hybrid analog/digital precoding/combining for hybrid architectures. Here, some of the signal processing is done analogically and another is done digitally. At transmitter side, in order to do an efficient beamforming processing, it is necessary to know the channel before the transmission known as Channel State Information (CSI). Considering frequency division duplex (FDD) mode, the CSI is estimated at the receiver and then sent to the transmitter via a feedback link. FDD mode requires two separate communication channels (frequency bands) and therefore we do not have channel reciprocity as in the time division duplex (TDD) mode, where channel information is estimated in the uplink. TDD uses a single frequency band for both transmit and receive. The CSI knowledge is needed to compute the hybrid analog-digital beamforming. However, perfect channel state information at the transmitter (CSIT), mainly for FDD systems, is not realistic in many practical scenarios.

An easily adaptable codebook design often used for MIMO communications is to randomly generate the codebook using Random Vector Quantization (RVQ) [19]. RVQ is a simple approach to codebook design that generates the vectors independently from a uniform distribution on the complex unit sphere. In a RVQ, the beamforming vector is restricted to lie in a codebook that is both known at the transmitter and receiver and it is randomly generated every time that the channel changes. The receiver chooses through its channel estimation one vector that maximizes the signal-to-noise ratio (SNR). The vector index is then sent to the transmitter. The main limitation of the RVQ is the

complexity since it requires very large codebooks, mainly for a large number of antennas as for the case of massive MIMO.

As we are facing systems with hundreds, or even thousands of antennas, the amount of CSI information needed to be sent by the feedback link would be huge and impracticable in real systems. Thus, a possible solution to overcome this problem is to send only a portion of the channel information, thereby designing the beamforming with a limited number of feedback bits sent from receiver to transmitter.

Thus, the aim of this dissertation is the development and the evaluation of the channel quantization strategies for massive MIMO mmWave, more efficient than RVQ. It is considered a transmitter that uses a hybrid analog/digital precoding/beamforming and a receiver that employs a hybrid analog/digital combining/equalizer specifically designed for this work. A simple yet efficient quantization strategy based on uniform quantization is proposed which consists in the quantization of some parameters of the mmWave channel, such as the fading coefficients and spatial angles.

1.4 Contributions

The main contribution of this dissertation is the development and evaluation of channel quantization schemes for broadband massive MIMO mmWave systems. Moreover, a hybrid analog-digital equalizer was also designed. The main work of this dissertation originated the following publication:

- S. Teodoro, A. Lopes, R. Magueta, A. Castanheira, A. Silva, R. Dinis. and A. Gameiro, "Performance Evaluation of a Frequency Selective Millimeter Wave System with Limited Feedback", *in proc. of International Congress on Ultra Modern Telecommunications and Control Systems* (ICUMT), Munich, Germany, Nov. 2017.

1.5 Structure

This dissertation is organized in 6 Chapters:

In Chapter 2 the MIMO systems are approached and explained in detail. Diversity and multiplexing characteristics are addressed.

Then, in the Chapter 3, a briefly summary of OFDM modulation is described. The basis and principles of quantization are also explained. It will be approached two different cases of quantization: Uniform Quantization and Random Vector Quantization.

In Chapter 4, the characteristics of the two 5G key technologies – mmWave and massive MIMO - are explained in detail and then, the combination of these two

technologies is studied as one of the most promising advances in 5G. Hybrid architectures are approached and then a common used mmWave channel model is presented.

In Chapter 5, it is proposed and evaluated an efficient limited feedback strategy for a hybrid mmWave massive MIMO OFDM system, where only a part of the parameters associated to the complex link channel are quantized and fed back. Firstly, it will be described in detail the hybrid mmWave massive MIMO OFDM system model. Secondly, a low-overhead uniform-based quantization strategy is described in detail, where only some CSI parameters are quantized. Finally, the system is evaluated under the proposed CSI quantization strategy and compared to the case where the perfect CSIT is known.

At the end, in Chapter 6, the conclusions of this work and some guidelines for future work and investigation are presented.

1.6 Notation

In this dissertation, the notation will be:

Boldface capital letters denote matrices and boldface lowercase letters denote column vectors. The operation $(.)^H$ represents the Hermitian transpose, of a matrix. For a complex number c, $R_e(c)$ and $\mathrm{Im}(c)$ represents the real part of c (imaginary part of c). $\mathbf{A}(j,l)$ denotes the element at row j and column l of the matrix $\mathbf{A}$. $\mathbf{I}_N$ is the identity matrix with size $N\mathrm{x}N$. The symbol * indicates that the transmitter should change the sign of quadrature component, in the process of complex conjugation.

Chapter 2

Multiple Antenna Systems

In the last few years, the exponential growth of the services on the wireless communication caused a huge impact in the telecommunications models and networks developments around the globe. It's widely expected that the number of connections continues this exponential growth. Along with this increase in the number of connection, the demand of consumers of wireless communication grows too, which leads to more services, higher data rates, more quality of service.

Multiple antennas are an important mean to improve the performance of wireless systems. It is widely understood in a system with multiple transmit and receive antennas (MIMO channel), the spectral efficiency is much higher than that of conventional single antenna channels. Research on multiple antenna channels [20], [21] and the design of communication schemes [22]- [23] demonstrates a great improvement of performance [24]. So, the use of multiple antennas techniques has been one of the major factors for the wireless systems evolution of, where the spatial dimension is achieved, creating several paths between the transmitter and the receiver [25].

In this chapter, various multiple antenna communication techniques are approached and discussed and how they can exploit the spatial dimension through diversity and spatial multiplexing.

2.1 MIMO System Definition

Traditionally, multiple antennas have been used to increase diversity to combat channel fading. Each set of transmit and receive antennas provide a signal path from the transmitter to the receiver. By sending signals with the same information through different paths, can be obtained in the receiver side multiple independently faded replicas of the data symbol required. More reliable reception is achieved [24].

The Single Input Single Output (SISO) is the simplest configuration, when transmitter and receiver only have one antenna each, as we can see in Fig. 2.1. Of course, the great advantage of this scheme is the simplicity because when both transmitter and receiver have only one antenna there is no diversity, so there isn't additional processing required. However, the SISO channel is limited in its performance. Interference and fading will have a seriously impact in the system more than a MIMO system using some form of diversity [26].

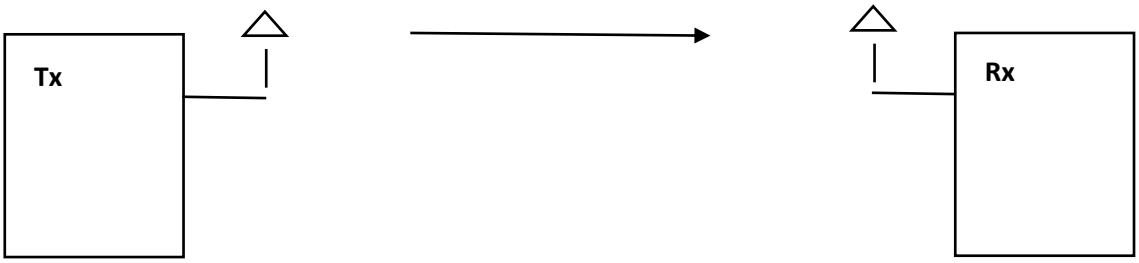

Figure 2.1 - SISO configuration

In Fig. 2.2 is represented a Single Input Multiple Output (SIMO) system. This version of MIMO occurs when the transmitter has a single antenna but the receiver got multiple ones. This is also known as receive diversity because of the multiple signals from several independent sources normally combat the effects of fading. Consequently, this requires a higher processing techniques on the receiver that can be a limitation because of the size, cost and battery drain in the case of mobile devices. Although, it is a configuration relatively simple to implement and has been used for many years with short listening/ receiving station to combat the effects of the ionospheric fading and interference [26].

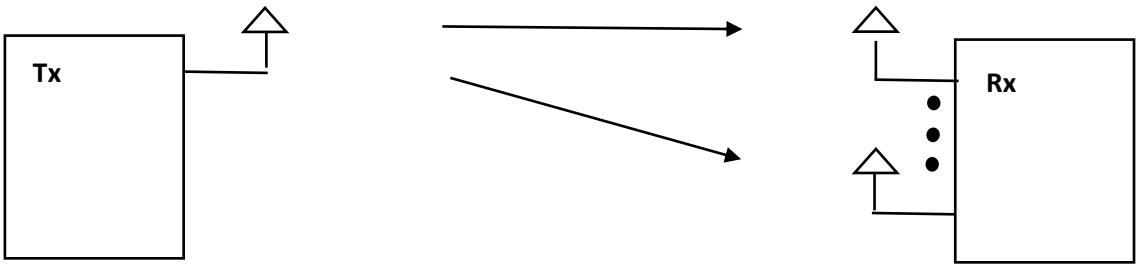

Figure 2.2 - SIMO configuration

Another configuration is illustrated in Fig. 2.3, the Multiple Input Single Output (MISO). MISO is when the transmitter is equipped with multiple antennas but only one in the receiver. This technique is also known as transmit diversity because the transmitter send the data in a redundantly way and then, the receiver has the capacity to receive the optimum signal and then extract the required data. The receiver can choose the best signal received. Now, the processing complexity is in the transmitter side, which is an advantage because, in the case of mobile phones, there is no much space to put a large number of antennas and the processing required on the receiver can be minor. This affects the size, cost and battery life of mobile phones [26].

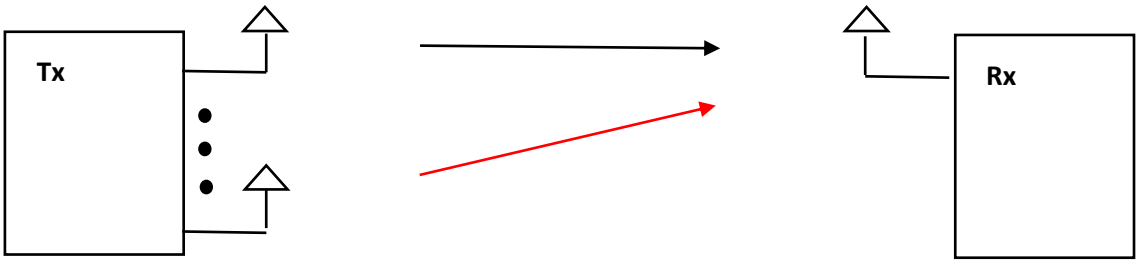

Figure 2.3- MISO configuration

The last configuration, illustrated in Fig. 2.4, is the MIMO scheme. in this configuration there is more than one antenna in each terminal, transmitter and receiver. Despite of the required complexity in both terminals, the improvements in the channel robustness and in the channel throughput overweight that disadvantage [26].

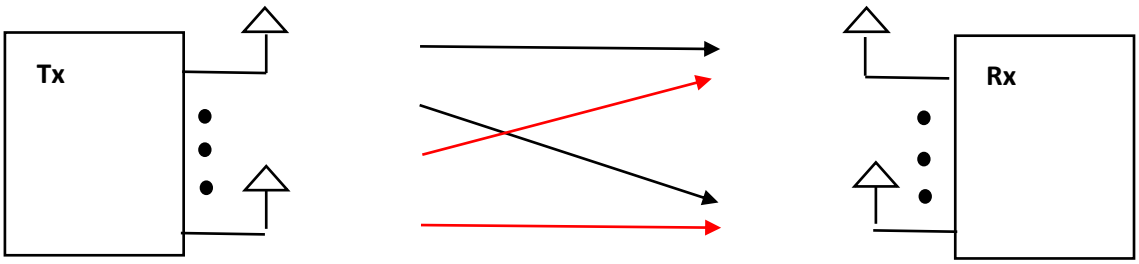

Figure 2.4 - MIMO configuration

2.1.1 Diversity

In MIMO systems, we must increase the reliability of the communication operation between transmitter and receiver, while maintaining a high spectral efficiency. The ultimate solution relies on in the use of diversity, which can be seen as some kind of redundancy [27].

If a radio signal is transmitted/ received through only one channel, then, a deep fade may cause the disappearing of the signal. The key is to create multi channels that have independent and uncorrelated fading. Diversity is mainly used in radio communications and is a common technique for combating fading and interference, and so avoiding error bursts. It is based on the fact that individual channels experience different levels of fading and interference. Multiple versions of the same signal may be transmitted or received. So, basically, diversity means that the same information flows through different independent channels [28] [29] [30]. Therefore, by applying these methods, the probability of all the replicas are affected by attenuation and noise are reduced. Diversity techniques may exploit the multipath propagation, resulting in a diversity gain [31], [32].There are various techniques used, including the next three:

- **Time diversity**

  In timer diversity, the same information is transmitted in different time slots. In practice, time diversity is usually associated with interleaving and coding over symbols across different coherent time periods. A good gain can be achieved when the duration between the two-adjacent slots - in which the same symbol is transmitted - is greater than the coherence time of the channel [27].

- **Frequency diversity**

  In the frequency diversity, the same information signal is transmitted on different subcarriers. The maximum diversity gain can be achieved when the separation between subcarriers is greater than the coherence bandwidth.

- **Spatial/ Antenna diversity**

  It is a method where more than one antenna is used to overcome the detrimental effects of multipath fading. In spatial diversity, the same information signal is transmitted or received via different antennas. The maximum gain can be achieved when the fading occurring in the channel is independent and uncorrelated. In the receiver, diversity gain can be achieved by combining the redundant signals arriving via independent and uncorrelated channels [27].

The single antenna diversity techniques – time and frequency diversity – have some drawbacks. Time diversity decreases the data rate by a factor of L, where the L is the number of independent paths where the information flows. In the case of the frequency diversity, this technique requires more bandwidth. On the other hand, spatial diversity is achieved without increasing the bandwidth.

In conventional wireless communications systems, spectral and power efficiency are achieved by exploiting time and frequency diversity techniques. However, the spatial dimension - so far only exploited for cell sectorization - will play a more important role in future wireless communication systems [33].

2.1.1.1 Receive Diversity

This technique, often associated to the SIMO architecture, occurs when the receiver has multiple antennas.

As we can see in Fig. 2.5, both signals are subject to fading and if that happen at the same time of the signal, then, the power combined at the receiver will be low, causing critical failures in the signal transmitted. To prevent this, if the receiver antennas are far enough apart from each others, then, the two sets of fading geometries will be very different, so, is far more likely that the signals will suffer fading at completely different times. It is very important that antennas must be wide enough – a few wavelengths from the frequency carrier - to prevent the mutual correlation between channels. The signals reach the multiple receive antennas with different phase shifts that can be removed in the antennas by channel estimation, thus, making possible the addition of the two signals in phase without any risk of destructive interference between them. Normally, BSs can have more than one receive antenna, so, receive diversity is more often used in the uplink. However, nowadays, in LTE, the mobile's specifications assume that the mobile is using two receive antennas, so, LTE systems are expected to use receive diversity in the downlink as well as the uplink. In mobile devices, the receiver antennas are closer than in the BS's, which reduces the benefits of the receive diversity [8].

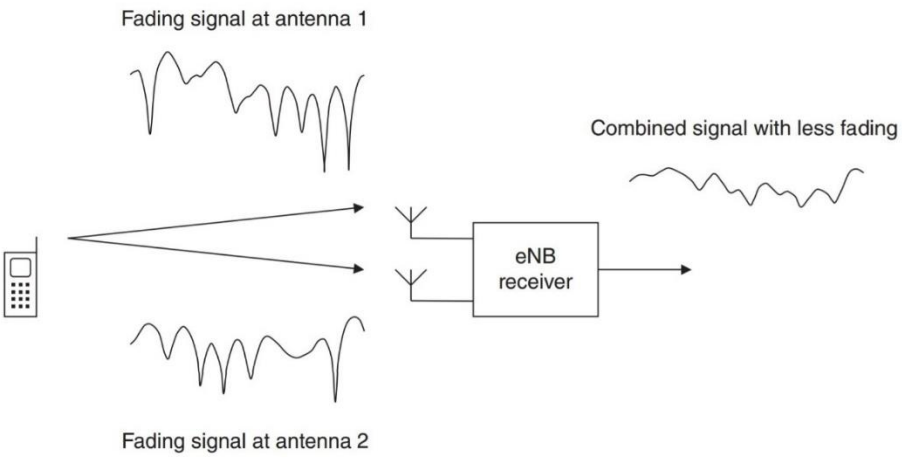

Figure 2.5 - Receive diversity: reduction in fading by the addition of two uncorrelated signals [8]

Receive diversity can achieve both diversity gain and antenna gain. The diversity gain is associated to the fact that the channels are independent and the antenna gain is related to the fact that the noise terms added at each receiver are independent [34].

The system performance evaluation is done by the combining of the multiple replicas sent by the transmitters, in order to increase the overall SNR. Assuming a architecture with M_t transmitter antennas and M_r receiver antennas, the use of M_r antennas allows the reception of the multiple transmissions sent by $M_r M_t$ number of channels and, by different types of processing/combining at the receiver, it is possible to improve the SNR.

Diversity scheme can also be classified according to the type of combing methods employed at the receiver. According to the implementation complexity and the level of channel state information required by the combining method at the receiver, there are four main types of combining techniques: Selection Combining (SC), Switched Combining (SWC), Equal Gain Combining (EGC) and Maximal Ratio Combining (MRC) [35] [36] [34].

Selection Combining

With this technique, the antenna with the stronger signal is used. SC selects as an output the largest instantaneous SNR at every symbol interval among the M_r received signal, so that the output SNR is equal to that of the best incoming signal.

The selection algorithm compare the instantaneous amplitude of each channel and choose a branch with the largest amplitude, thus the received signals from others antennas are ignored. In practice, the signal with the highest sum of the signal and noise power is usually used, since it is difficult to measure SNR [35].Antenna gain increase with the number of branches, but not linearly. Increasing M_r yields diminishing return in terms of antenna gain. The main antenna gain happens when $M_r = 1$ and $M_r = 2$ .The SC achieve the same diversity gain as MRC and EGC, i.e., M_r .

This technique can be implemented with just one antenna receiver, which is clearly desirable from a cost, size and power consumption view and this scheme does not require de knowledge of the channel state information.

Switched Combining

In this combining technique, the receivers scan all the diversity branches and choose one branch with a SNR above than a certain predetermined threshold. This signal is selected as the output. If it SNR drops below the threshold line the signal is no longer selected to the output and the receiver start scanning again among all the diversity branches and will switch to another one. This does not improve the performance, but it

does improve the likelihood. This is also known as scanning diversity or threshold diversity.

Compared to SC, the SWC is inferior because it does not continually search for a better instantaneous signal. However, since it does not require a continually and simultaneously monitoring, switched combining is simpler to implement. Like in the SC scheme, this technique uses as an output signal only one of all the diversity branches and does not require the knowledge of the channel state information. SC and SWC can be used in conjunction.

Equal Gain Combining

This scheme is a suboptimal but simpler linear combining method where the estimation of the amplitude fading for each individual branch is not required. Instead, the receiver sets the amplitudes of the weighting factor to be unitary. Here, the signals from antennas are co-phased prior to summation, i.e., the EGC co-phases the signal on each branch and then combines them equal weighting. EGC only performs phase rotation on the received signals. Antenna gain also increases linearly with the number of branches but it is lower than the MRC technique. The performance of EGC is only marginally inferior to MRC, since the antenna gain is lower. However, the implementation complexity for EGC is significantly less than the MRC. By using EGC technique is required two receiver antennas.

Maximal Ratio Combining

MRC is a linear combining technique that various signals inputs are individually weighted and added together to get an output signal. Each individual signal is co-phased and summed with optimal weighting – weighted with its corresponding amplitude and signal phases - to maximize the combiner output SNR. The maximum output SNR is equal to the sum of instantaneous SNRs of the individual signals. This scheme is also known as matched filter or optimum combining since it can maximize the output SNR.

Like said above, receive diversity can achieve both space diversity gain and antenna gain. In this case the space diversity gain is equal to the number of receiver branches (M_r), assuming independent and identical distributed Rayleigh channels. The antenna gain increase linearly with the number of branches.

In the uplink, the MRC is the most promising single-user detection technique since the spreading codes do not superpose in an orthogonal fashion at the receiver and maximization of the SNR ratio is optimized. This scheme requires the knowledge of channel fading amplitude and signal phases.

2.1.1.2 Transmit Diversity

For many reasons, multiple antennas can be installed at just one link end, usually the BS. For the uplink transmission, from the mobile station (MS) to the BS, multiple antennas can act as receive diversity branches, compensating for the relatively low transmission power from the mobile, thus improving the quality and the range of the uplink. On the other hand, for the downlink, it is difficult to utilize receive diversity at the mobile. It is hard to compact a large number of antennas in a small mobile equipment and multiple receive antennas imply multiple sets of RF convertors and, as a result, more processing power, which is limited for mobile units. So, for downlink, is more suitable the transmit diversity [35]. There are some ways of transmitting signals from several transmitters/ antennas and achieve diversity effects with it. However, at the receiver side, can be difficult to reconstruct the signals sent by multiple antennas. Signals with interference can be added in a destructive way. To overcome this problem a precoder or a space-time/frequency coding can be used. There are two kind of techniques: close loop and open loop.

Close Loop

In the close loop technique is taken into account the Channel State Information (CSI) to improve the performance of the system. In this technique, the transmitter send the signal but applies a phase shift to one or both copies of the signal. The phase shift is necessary to avoid the destructive interference at the receiver. To ensure that the signals are received on phase, the receiver fed back to the transmitter as a Precoding Matrix Indicator (PMI). This precoder teach the transmitter if the signal must be phase shifted or not.

However, this technique is only suitable for slowly moving mobiles. In practical cellular mobile system, mobility and environment change cause fast channel variations, making channel tracking difficult. If the mobile is moving fast, the PMI that is feedback will change and cannot correspond to the best choice when it is received by the transmitter, because the CSI changes very quickly [8]. The imperfect channel estimation and mismatch between previous channel state and current channel condition will decrease the received signal SNR and affect the system performance.

Open Loop

In the open loop scheme, the CSI knowledge is not required thus allowing a greater robustness under unfavorable conditions. Signal processing at the transmitter is designed properly to enable the receiver exploiting the embedded diversity from the received signals [35]. To achieve better performances in multiple antenna transmission, it is possible to combine error control coding, modulation and transmit diversity design with no bandwidth expansion. Coding technique designs for multiple antenna transmission are called space-time/frequency coding. These codes add in a proper way redundancy in both spatial and temporal domains, which introduces correlation into transmitted signals. This is of course of particular importance on the downlink of mobile radio systems, since it is easier to provide multiple antennas in a BS than in a mobile handset. The principle of these codes is to provide through coding constructive superposition of the signals transmitted from different antennas [33]. These codes can be further combined with multiple received antennas to minimize the effects of multipath fading and to achieve the capacity of massive MIMO systems.

There are mainly two types of space-time codes: trellis and block codes. Tarokh formalized the concept of a space-time code, describing the criteria we have considered above, and introduced the space-time trellis codes. At the same time, Alamouti described what was in fact the first space-time block codes. Space-time trellis codes to achieve full diversity, for a large number of antennas, require to the encoder to have longer memory, which can be excessively computationally expensive [37]. So, the first ones normally outperform the block codes but with a higher complexity inherent and, for that reason, their use in practical systems are limited. Despite this performance penalty, Alamouti scheme is still appealing in terms of simplicity and performance. There are two types of block codes: space-time block codes (STBC) and space-frequency block codes (SFBC).

Fig 2.6 shows an implementation of an open loop transmit diversity technique that is known as Alamouti coding. This is the simplest coding scheme with no bandwidth extension and uses orthogonal codes for two transmit antennas and M_r receiver antennas. This technique was adopted in the LTE standards. Here, the transmitter will send two symbols (s_1 and s_2), in two consecutive time steps. In the first step, the transmitter, will send s_1 from the first antenna and s_2 from the second antenna. In the second-time step, the first antenna will send the symbol $-s_2^*$ and the second antenna will send the symbol s_1^*.

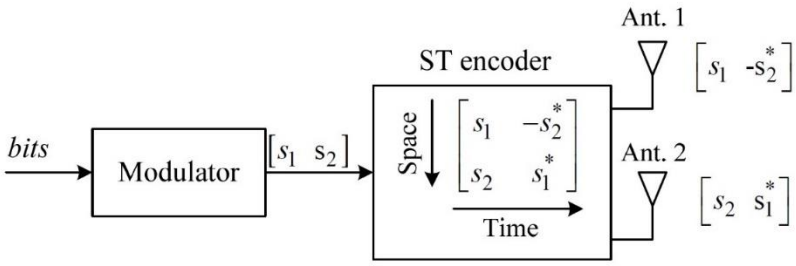

Figure 2.6 – Operation in the transmitter of Alamouti scheme for open loop transmit [34]

The matrix $\mathbf{S}$ resumes this process:

$$\mathbf{S} = \begin{bmatrix} s_1 & -s_2^* \\ s_2 & s_1^* \end{bmatrix}. \tag{2.1}$$

Fig. 2.7 describes the receiver side of the Alamouti coding; $\mathbf{h} = [h_1 \quad h_2]$ is the vector channel, being h_1 and h_2 the coefficients of antennas 1 and 2, respectively; $\mathbf{n} = [n_1 \quad n_2]$ is the vector from additive white Gaussian noise; $\mathbf{y}$ is the received signal matrix.

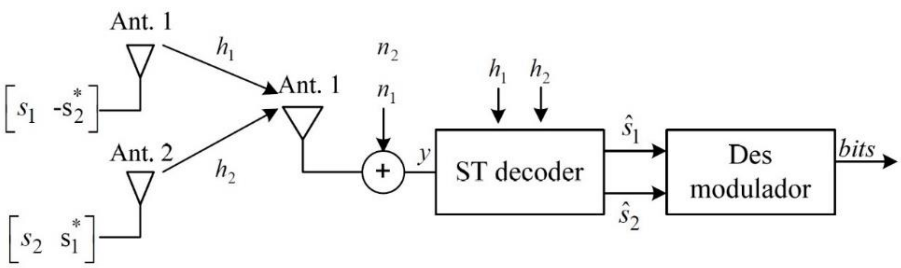

Figure 2.7 – Operation in the receiver of Alamouti scheme for open loop transmit [34]

$$\mathbf{y} = \frac{1}{\sqrt{2}}\mathbf{hS} + \mathbf{n}. \tag{2.2}$$

The $\frac{1}{\sqrt{2}}$ factor is used in order to normalize the power of each symbol to 1. The previous equation can be written as:

$$\begin{cases} y_1 = \frac{1}{\sqrt{2}} h_{1,n_1} s_1 + \frac{1}{\sqrt{2}} h_{2,n_1} s_2 + n_1 \\ y_2 = -\frac{1}{\sqrt{2}} h_{1,n_2} s_2^* + \frac{1}{\sqrt{2}} h_{2,n_2} s_1^* + n_2 \end{cases}. \tag{2.3}$$

After the Alamouti decoding, the soft decision is given by:

$$\begin{cases} \hat{s}_1 = \frac{1}{\sqrt{2}} h_{1,n_2}^* y_1 + \frac{1}{\sqrt{2}} h_{2,n_1} y_2^* \\ \hat{s}_2 = \frac{1}{\sqrt{2}} h_{2,n_1}^* y_1 + \frac{1}{\sqrt{2}} h_{1,n_1} y_2^* \end{cases}. \tag{2.4}$$

Assuming $h_{k,n} = h_{k,n+1}$, i.e., the channel between two adjacent frequencies are extremely correlated, so, the expression gets simplified:

$$\hat{s}_n = \frac{1}{2}\left(\left|h_{1,n_1}\right|^2 + \left|h_{2,n_1}\right|^2\right) s_n + \frac{1}{\sqrt{2}} h_{1,n_1}^* n_1 + \frac{1}{\sqrt{2}} h_{2,n_1} n_2^*. \tag{2.5}$$

The SNR is given by:

$$SNR = \frac{1}{2}\left(\frac{\left|h_1\right|^2 + \left|h_2\right|^2}{\sigma^2}\right). \tag{2.6}$$

The previous analysis was made for a system with two transmitter antennas but only one in the receiver side, where Alamouti scheme can achieve a diversity of order 2. The previous analysis can be generalized to more than one antenna in the receiver side, expecting then a $2M_r$ diversity order. Although, the Alamouti STBC scheme approached here can only be available with 2 antennas maximum in the transmitter.

The paradigm approached above make the use of the orthogonal designs to design space time block codes. While providing the full diversity, the codes presented also provide the maximum possible transmission rate allowed by the theory for real signal constellations. However, a complex orthogonal design and the corresponding space time block code which provide full diversity and full transmission rate is not possible for more than two transmitter antennas [38]. A good solution to this problem are the Tarokh orthogonal codes or the Quasi-orthogonal codes. Tarokh codes are orthogonal like the Alamouti codes and has the advantage that they can be used when there are more than two antennas in the transmitter. However, in spite of these advantages, Tarokh codes have the disadvantage of having a code rate less than one, and because of that, it is required a bandwidth expansion. Because of the lower rate, the Tarokh codes can achieve full orthogonality between streams in each antenna, making full diversity order possible. To solve these problems, Quasi-orthogonal codes were proposed which have a code rate of one, but cannot achieve full diversity.

2.1.2 Spatial Multiplexing

Spatial multiplexing has a different purpose from diversity processing [8]. SIMO and MISO systems provide diversity and antenna gain but no multiplexing gain, also called degree of freedom (DoF). If the system is constituted by multiple antennas at both transmitter and receiver, then we can configure multiple parallel data streams between them, to increase the data rate. In a system with M_t transmitter and M_r receiver antennas, often known as $M_t \mathrm{x} M_r$ spatial multiplexing system, the peak data rate is proportional to min (M_t, M_r). This data rate can be increased for the same bandwidth and with no additional power. This is only possible in MIMO systems. With multiple parallel data streams, to achieve high spatial gain, low channel correlation is required. A high degree of difference between channels is needed to perform the separation of multiple layers without interference between them. MIMO spatial multiplexing also enable improvement in cell capacity and throughput [39].

However, even if we can combine multiplexing gain and diversity gain in the same system, the fulfillment of the two gains is impossible to achieve. For example, in a 2 x 2 system, to have full diversity, each symbol must go through two independent channel and so diversity achieved is $d = 2$ and multiplexing gain is $r = 1$. In the other hand, to fulfill the multiplexing gain at the same system, each channel is used by only one data stream, $r = 2$, but there is no repetition of the symbols, so the diversity is $d = 1$. It is clear that multiple antennas at the receiver and at transmitter can be used to improve the receiver SNR ratio in proportion to the number of antennas. In both gains, there is a simple relation between SNR and the bit error rate (BER). Like said in the section above, in the diversity, the data rate is constant and the BER decreases as the SNR increases. In multiplexing, the BER is constant while the data rate increases with the SNR [37].

2.1.2.1 Single User MIMO Techniques

Chanel Known at the Transmitter

The optimal way of communicating over MIMO channels involves a channel-dependent precoder, which fulfils the roles of both transmit beamforming and power allocation across the transmitted streams, and a matching receive beamforming structure. Full channel knowledge is therefore required at the transmit side for this method to be applicable.

Let us consider the system model represented in the Fig. 2.8, when CSI is known at the transmitter. In the transmitter, the data symbols are precoded into different

antennas. The precoders are computed based on the channel information that should be feedback from user's equipment to the BSs. [34]. Here the $\mathbf{W}$ is the precoding/ beamforming matrix and $\mathbf{G}$ is the equalization or post-processing matrix.

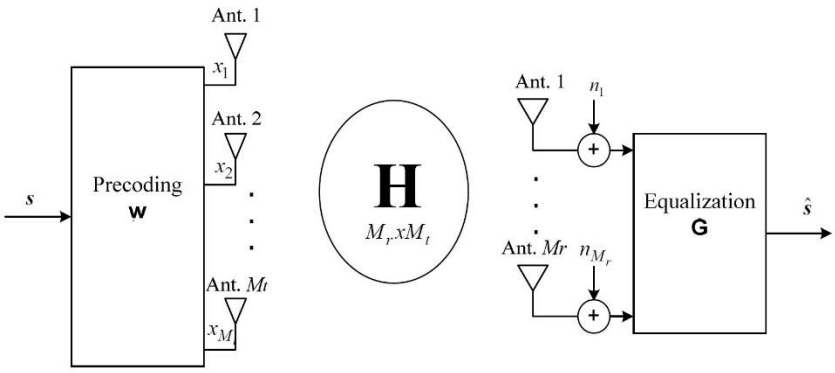

Figure 2.8 - MIMO scheme for spatial multiplexing system

In matrix notation, we can write the received signal as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \,. \tag{2.7}$$

Here, $\mathbf{x}$ is a column vector that contains the signals that are sent from the M_t transmit antennas. Similarly, $\mathbf{n}$ and $\mathbf{y}$ are column vectors containing the noise and the resulting signals at the M_r receive antenna. The channel matrix $\mathbf{H}$ expresses the amplitude changes and phase shifts that the air interface introduced. The matrix has M_r rows and M_t columns, so it is knowns the $M_r \mathrm{x} M_t$ matrix.

$$\mathbf{H} = \begin{bmatrix} h_{1,1} & \cdots & h_{1,j} \\ \vdots & \ddots & \vdots \\ h_{i,1} & \cdots & h_{i,j} \end{bmatrix}, \mathrm{i} = 1,\ldots,\mathrm{M}_r;\ j = 1,\ldots,M_t \,. \tag{2.8}$$

$$\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_{M_t} \end{bmatrix}^T . \tag{2.9}$$

$$\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_{M_r} \end{bmatrix}^T . \tag{2.10}$$

$$\mathbf{n} = \begin{bmatrix} n_1 & \cdots & n_{M_t} \end{bmatrix}^T . \tag{2.11}$$

The MIMO channel can be converted into a set of parallel channels. This can be done by decomposing the $\mathbf{H}$ matrix, using a singular value decomposition (SVD). With SVD we are able to estimate the capacity of each channel in order to select the best channel to adapt the transmission. This adaptation is done by performing a power allocation according the singular values computed using the SVD technique. We also obtain the

optimum signal precoding to perform at the receiver. Denoting that the CSI must be available at both transmitter and receiver sides [40]. So, using SVD, the $\mathbf{H}$ matrix is now given by:

$$\mathbf{H} = \mathbf{UDV}^H . \tag{2.12}$$

Here, the $\mathbf{U}$ and the $\mathbf{V}$ are unitary matrices of size $M_r \times r$ and $M_t \times r$, respectively. r is a rank indication resulting from the measure, by the receiver of the system, of the channel elements. It indicates how much symbols the system can successfully receive.

$$r = rank(\mathbf{H}) \leq \min(M_t, M_r). \tag{2.13}$$

The matrix $\mathbf{D}$ is a diagonal matrix whose diagonal elements are non-negative real numbers:

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r \end{bmatrix}. \tag{2.14}$$

The precoder $\mathbf{W}$, with the size of $M_t \times r$, is set as:

$$\mathbf{W} = \mathbf{VP}^{\frac{1}{2}} . \tag{2.15}$$

where, the $\mathbf{P}$ is a square diagonal power allocation matrix of size $r \times r$:

$$P = \begin{bmatrix} p_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_r \end{bmatrix}^{\frac{1}{2}}. \tag{2.16}$$

The equalizer matrix $\mathbf{G}$ is given by:

$$\mathbf{G} = \mathbf{U}^H . \tag{2.17}$$

The transmitted signal $\mathbf{x}$ is given by:

$$\mathbf{x} = \mathbf{Ws} . \tag{2.18}$$

$\mathbf{W}$ is the precoder matrix and $\mathbf{s}$ is the data stream transmitted over the M_t antennas; $\mathbf{s}$ is the data vector of the size $r \times 1$:

$$\mathbf{s} = \begin{bmatrix} s_1 & \cdots & s_r \end{bmatrix}^T . \tag{2.19}$$

If we replace $\mathbf{W}$ and $\mathbf{H} = \mathbf{UDV}^H$ on the received signal by:

$$\mathbf{y} = \mathbf{UDV}^H \mathbf{VP}^{\frac{1}{2}} \mathbf{s} + \mathbf{n} . \tag{2.20}$$

The estimated transmitted symbol is obtained by:

$$\hat{\mathbf{s}} = \mathbf{Gy} = \mathbf{U}^H \mathbf{UDV}^H \mathbf{VP}^{\frac{1}{2}} \mathbf{s} + \mathbf{U}^H \mathbf{n} \,. \tag{2.21}$$

$$\hat{\mathbf{s}} = \mathbf{DP}^{\frac{1}{2}} \mathbf{s} + \hat{\mathbf{n}} \,. \tag{2.22}$$

The soft estimated of the rth symbol is:

$$\hat{s}_i = \lambda_i \sqrt{p_i s_i} + \hat{n}_i, i = 1, ..., r \,. \tag{2.23}$$

Here λ represents the amplitudes of the channel. If the channel is known in the transmitter, converting the channel MIMO into r parallel channels through SVD it is possible to transmit r parallel free interference data symbols.

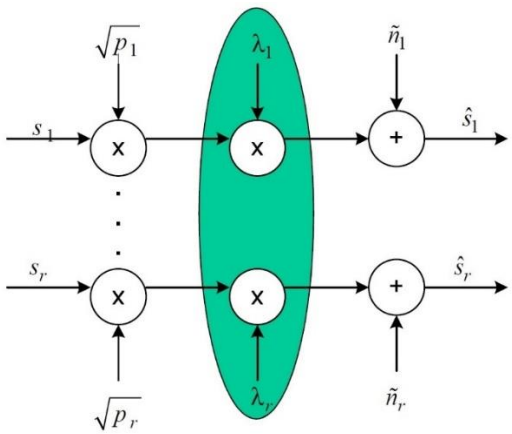

Figure 2.9 - The symbols estimation scheme

One important aspect is to constraint the transmit power to a given value P_t, i.e.:

$$\sum_{i=1}^{r} p_i = P_t \,. \tag{2.24}$$

When the CSI is known at the transmitter, depending on the metric we want to maximize, the best strategy to allocate the available power for the different data symbols is the Waterfilling. This technique is used to maximize the capacity using a limit to decide the amount of power allocated in each channel. If the SNR of a channel is so low that result in water level override, the power allocated in this channel will be 0, transmitting least 1 symbol in the same frequency [41]. So, the channel is discarded, and the available power will be distributed by the “good” channels. At high values of SNR, the performance is the equivalent. But, at lower values of SNR, the Waterfilling allocation provides higher gains.

Chanel not Known at the transmitter

It was shown that with the CSI at the transmitter, the capacity of the system balance linearly with $rank(\mathbf{H}) \leq \min(M_t, M_r)$. In that case, capacity is slightly higher due to the Waterfilling power allocation. However, sometimes this kind of knowledge is not possible at the transmitter, so there are no conditions to perform the SVD decomposition technique. Without the CSI knowledge at the transmitter the choice of the coordinate system in which the independent data streams are multiplexed has to be fixed a priori. In conjunction with joint decoding, we see that this transmitter architecture achieves the capacity of the fast fading channel. This architecture is also known as Vertical-Bell Labs Space-Time Architecture (V-Blast) [29]. This is the simplest possible layered space-time structure [42].

To summarize, with this architecture that independent data streams are firstly demultiplexed in the transmitter into M_t parallel streams, each of which is encoded separately. Each encoded data stream is then transmitted from a different antenna. The channel mixes up the different data streams and then, the receiver separates them out by successive nulling and interference subtraction. Note that the encoding scheme does not need the cooperation between different antenna elements nor users. This technique uses the first data stream as the useful one and regard the other ones as interference. Because of that, V-Blast faces the problem of error propagation, i.e., if the data stream 1 is decode incorrectly, we subtract the wrong signal from the remaining signals at the antenna elements.

However, the main drawback of the V-Blast architecture is that complexity grows exponentially with the number of data streams and that is very impractical to achieve the system efficiency. To solve this problem there are linear sub-optimal receiver architectures. These linear equalization schemes performed at the receiver like Zero-Forcing (ZF) based equalizer, Minimum Mean Square Error (MMSE) based equalizer and Interference Cancellation techniques like Successive Interference Cancelation (SIC) technique. The system architecture used in linear equalizers is known as Diagonal – Bell Labs Space-Time Architecture (D-Blast). With is technique the problem of the joint decoder is now solved. D-Blast decode the data streams each by each.

- **Zero-Forcing Equalizer**

  The aim of this technique is to design an equalizer vector for each data symbol that removes the interference. ZF applies channel inversion and can eliminate multiple access interference by restoring the orthogonality between the spread data [33]. All data symbols are detected without interference. However, the ZF technique has a drawback. At lowers SNR, this equalizer enhances noise, especially when the channel is in deep fading.

- **Minimum Mean Square Error Equalizer**

  The MMSE technique minimizes the mean square error between the transmitted symbol vector $\mathbf{s}$ and estimate $\hat{\mathbf{s}}$ at the receiver. The computation of the MMSE equalization coefficients requires the knowledge of the actual variance of the noise, σ^2. For SNR $\rightarrow\infty$, the MMSE equalizer becomes identical to ZF equalizer. To overcome the additional complexity for the estimation of σ^2, a low complex sub-optimal MMSE equalization can be performed [33].

- **Successive Interference Cancellation**

  This technique is like an addition to the previous ones. SIC can improve the performances of the of the linear ZF and MMSE equalizers, establishing a compromise between performance and complexity.
  Linear equalizer like ZF or MMSE are used to decode the first data symbol and then subtract off the symbol from the received vector $\mathbf{y}$. If the first data is correctly detected, then, the second equalizer only has to deal with data symbols as interference, since the first one has correctly subtracted. This process continuous to proceed until the final equalizer does not deal with any interference from the others data symbols.
  The main drawback of this technique is the error propagation since if the first data symbol is incorrectly detected, its error propagates to all the other streams further. To prevent this error to happen, one among solutions is to detect first the data symbols with higher average power and only at the end the one with smaller overall power [33].

2.1.2.2 Multi User MIMO Techniques

So far it has been approached the role of multiple transmit and receive antennas in the context of point-to-point channels. Point-to-point MIMO communications enable higher gains in both channel capacity and reliability, essentially via space-time codes and multiplexed transmissions symbioses.

However, the spatial dimension can be also used to separate the users that share the same frequency and time resources. This is a multiuser MIMO (MU MIMO). Such multiple access protocol implies an extra hardware cost, like antennas and filters, but does not involve any bandwidth expansion, unlike TDMA or CDMA [43]. There are two different types of MU-MIMO, such as the one used in downlink (the communication is done from the transmitter to different users/receivers; one-to-many) and the other one used in the uplink (the communication is done from different users/transmitters to the receiver; many to one). MU-MIMO is used more often in the uplink. Here, in the uplink,

the basic principles of MU-MIMO, like the receiver structures, are the same as the ones used in the Single User-MIMO (SU-MIMO), since each user terminal can be seen as one antenna in the previous point-to-point MIMO system. The downlink, in MU-MIMO, is the most challenging case. Like said above, the basic principles of MU-MIMO are very similar to the ones used in SU-MIMO, however, in the last subsections we saw that the equalization techniques like ZF or MMSE performed to separate the layers were done in the receiver side, while in the MU-MIMO this performance is executed in the transmitter side. The use of ZF or MMSE in the transmitter can only happen because the user terminals are probably sufficiently far from each other, which means that theirs signal paths are different. It is important to refer that the number of users simultaneously served is limited by the number of antennas at the receiver, i.e., the number of receiver antennas must be greater of equal to the number of users. For example, 1 receiver with 4 antennas can support 4 users each one with 1 antenna or 2 users each one with 2 antennas.

Moreover, in the SU-MIMO the processing in the signal is done when the signal is already affected by the noise. In MU-MIMO, the processing is done in the transmitter side, so, we anticipate the channel effects in the signal, and depending on this, the transmitted signal is adapted before the channels conditions affected it. In mathematical terms, there are no difference between remove the interference in the transmitter or in the receiver, but, if it is needed the anticipation of the channel behavior and conditions, the main challenges in the implementation of MU-MIMO is the need of CSI knowledge at the transmitter side to properly serve the spatially multiplexed users and also the complexity of the scheduling procedure associated with the selection of a group of users that will be served simultaneously. Optimal scheduling involves exhaustive search, whose complexity is exponential to the group size and depends on the precoding, decoding and the channel state feedback technique [43]. The uplink MU-MIMO block diagram is illustrated in Fig.2.10.

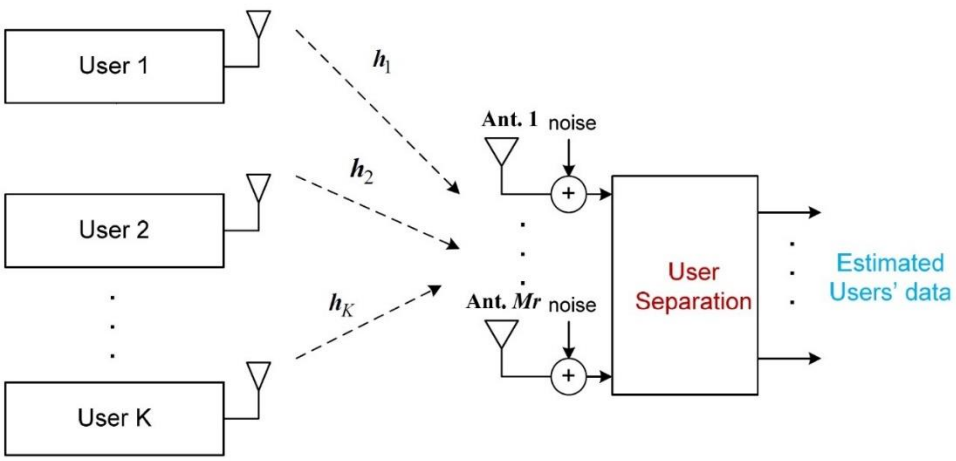

Figure 2.10 - MU-MIMO uplink block diagram

The uplink MU-MIMO, does not increase the peak data rates of each user mobile but enable higher cell throughput and can be implemented in non-expensive mobiles

with only one antenna [40].In the uplink, in spite of transmitters cannot cooperate, receivers can jointly process the received signal at all antennas. Like said above, despite the complexity associated to its implementation, is also possible using the MU-MIMO technologies for the downlink. In the downlink are the receivers that not cooperate between them. The signals are superimposed and interfere with each others, so, a precoder is required to mitigate the interference that a giver user can cause in other user terminals. It is important that the transmitter can do the tracking on the channel. If the transmitter does not track the channel, the precoding will not be performed in the downlink. MU-MIMO for downlink beamforming can generally use linear equalization like in the point-to-point MIMO systems. However, in this case, to use ZF or MMSE filters the number of transmitter antennas must be bigger or equal to the number of users [34].

As it says in [43], MU-MIMO techniques and performance have begun to be strongly developed because of the many key advantages over SU-MIMO communications:

- MU-MIMO can achieve direct gain in multiple access capacity proportional to the number of BSs due to the multiuser multiplexing schemes.
- MU-MIMO appears more immune to most propagation limitations comparing to SU-MIMO such as channel rank loss or antenna correlation. Moreover, the line of sight, which causes severe damage in SU-MIMO signals, is no longer an issue in multiuser schemes.
- MU-MIMO enable the spatial multiplexing gain at BS's without the need of multiple antenna terminals, thus, promoting this way the development of small and cheap terminals, without neglecting the relation between cost and intelligence.

Chapter 3

OFDM Modulation and Channel Quantization

The reasons that triggered the development and the emergence of new technologies for mobile communications were approached previously. Spectrum restrictions, increasing user data rates and demanding communication quality are some of the main operator problems in the future of telecommunications. Some of the techniques like multi antenna systems and multiuser communications discussed in the previous chapters are often used to achieve these demands, by providing communication links with substantial diversity and capacity and allowing the transmission of multiple spatially multiplexed data streams to multiple users, which results in very high data rates [20] [44] . Linear precoding, a generalization of beamforming at the transmitter used in this type of systems, aims to cancel the interference between users or separate the data symbols [45] [46]. However, to achieve efficiency in such systems and to orient the beam in space, the knowledge of the CSI in the transmitter side before any transmission is imperative.

For the particular case of massive MIMO based systems, this problem is more significant since terminals are equipped with a large number of antennas and therefore a huge amount of channels need to be fed back from the receiver to the transmitter, increasing the overall system signaling. In order to reduce this overhead some authors consider a strong spatial correlation, where only a few strong eigendirections need to be quantized or that the impulse response is sparse in time.

A technique that allows non-overloading of the feedback is the RVQ [19]. Random vector quantization is a simple approach of codebook design, firstly defined in [47], [48] where the vectors are generated independently from a uniform distribution on the complex unit sphere. Although RVQ techniques allows efficient limited feedback for multi antenna multi user schemes, the required RVQ's computational complexity and dimension codebooks can be very large, especially when we have a high number of transmitter and receiver antennas, thus making it very impractical for massive MIMO application [49]. In order to avoid the limitations imposed by the large RVQ codebooks and pursuing the reduction of the overhead without damaging the performance, a low complex limited feedback but efficient strategy based on Uniform Quantization (UQ) can be approached, where only some of parameters of the system channel are quantized.

3.1 Multicarrier Modulation: OFDM

The transmission of the information in wireless communications systems can be made by different techniques. Modulations are an essential part of the transmission. Usually, the main difference between single carrier and multicarrier modulations is the domain in which information is transmitted. Single carrier modulations are made in the time domain while in the multicarrier modulations the information is transmitted in frequency domain. The basic principle of multicarrier transmission is to convert a serial high rate data stream into multiple parallel low rate sub-streams. Since the symbol rate on each subcarrier is much less than the initial serial data symbol rate, the effects of inter symbol interference (ISI) significantly decrease, reducing the complexity of the equalizer [33].

One of the challenges in wireless systems is the severe frequency selective fading (FSF) caused due to the multipath channels between the transmitter and the receiver. The signal bandwidth in typical cellular systems exceed the coherence bandwidth of the multipath and consequently leads to ISI, which is usually dealt by the physical layer solutions like OFDM.

The OFDM transmission system technique has established itself as one elegant and popular method to overcoming the FSF in wireless systems. Some of the keys concepts in OFDM include the use of the orthogonality for sending several data symbols in parallel, cyclic prefix (CP) and simple equalization methods at receiver. Also, the use of multiple antennas to enhance the spectral efficiency and reliability are simpler using multicarrier systems like OFDM comparing to single carrier modulation schemes because the multicarrier nature transforms a broadband transmission into a multipath fading channel to several narrowband transmissions [50].

Important components in the OFDM schemes are the Inverse Fast Fourier Transform (IFFT) and Fast Fourier Transform (FFT) blocks in the transmitter and in the receiver, respectively. The samples of the transmitted OFDM signal can be obtained by performing an IFFT operation on the group of the data symbols that must be transmitted on orthogonal subcarriers. So, to recover the data, a FFT operation is performed in the receiver.

As can be seen in the Fig. 3.1, each subcarrier is overlapping the others, there is no mutual interference when the sampling is done at certain specific points in the frequency domain. This one of the reasons why higher spectral efficiency is achieved by the orthogonality between subcarriers as compared with other systems.

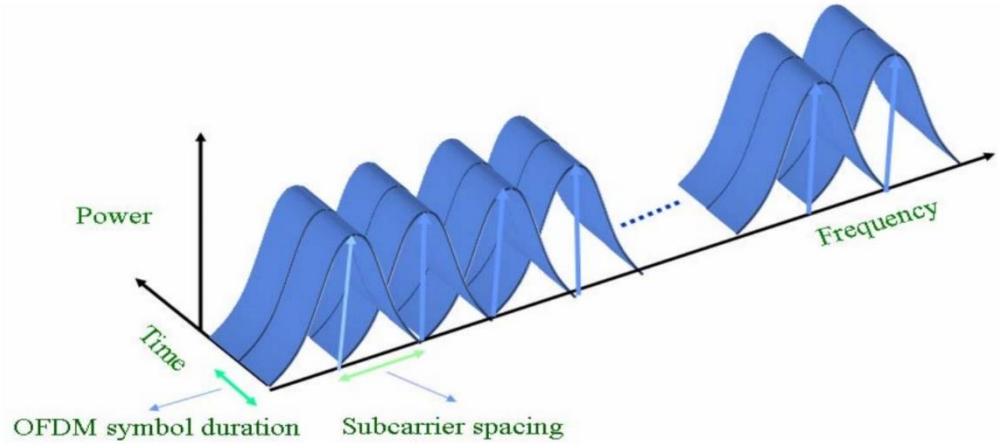

Figure 3.1 - Time-Frequency view of OFDM signal [50]

Each OFDM symbol has a duration of:

$$T_s = N_c T_d \,. \tag{3.1}$$

where T_s is the OFDM symbol duration, N_c is the number of subcarriers and T_d is the source symbol duration. The frequency spacing between two adjacent subcarriers is:

$$\Delta_f = \frac{1}{T_s} \,. \tag{3.2}$$

The frequency of each subcarrier:

$$f_n = n\Delta_f \,, n = 0, ..., N_c - 1 \,. \tag{3.3}$$

When the number of subcarriers increase, the OFDM symbol duration T_s becomes large and the amount of ISI reduces. However, to completely avoid the effects of ISI and thus to maintain the orthogonality between signals on the subcarriers to avoid ICI (Inter Carrier Interference) a guard interval of duration T_g is added. It is called cyclic prefix. This guard interval must be longer than the channel delay spread. If this happens, the receiver can be confident of reading information of just one symbol at a time, without any overlap with the symbols that precede or follow [51].

Consequently, the total duration of the transmitted OFDM symbol is:

$$T_s^{'} = T_g + T_s \,. \tag{3.4}$$

Illustrated in Fig. 3.2, CP is a copy or a duplication of the end part of the symbol and inserted in the start of it. Of course, this means increasing the full size of that symbol, but the gains overweight.

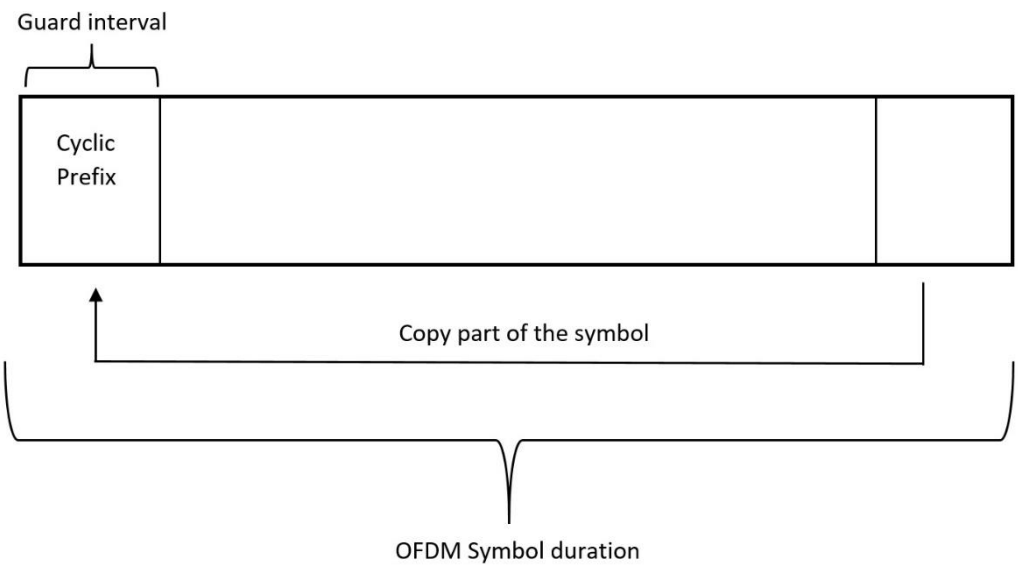

Figure 3.2 - The adding of the Cyclic Prefix

Nowadays, the deployed LTE technology uses a time guard, i.e., a CP of 4.7 µs. This corresponds to a maximum path difference of 1.4 Km between length of the longest and shortest rays, which is enough for all but the very largest and most cluttered cells. The CP reduces the data rate by about 7%, which is a small price to pay [51].

The Fig. 3.3 shows the digital multicarrier transmission system applying OFDM. At the entry, the data bits are going to be modulated into symbols. After this step the symbols changes from serial to parallel, which mixes each symbol with one of subcarriers by adjusting his phase and amplitude. After that, it is made a converse from frequency domain to time domain. This is made by an Inverse Discrete Fourier Transform (IDFT) or by a computationally more efficient IFFT. Then the signal is converted to serial again, CP is added to each OFDM symbol like seen above and after that a digital to analog conversion is made before transmission. At the receiver, the scheme sequence is identical but inverted. The signal is converted again to digital and CP is then removed. The conversion to parallel is made and then a Discrete Fourier Transform (DFT) or FFT is performed, changing from time domain to frequency domain. Again, a parallel to serial conversion is made and then the symbols are demodulated to obtain the transmitted signal.

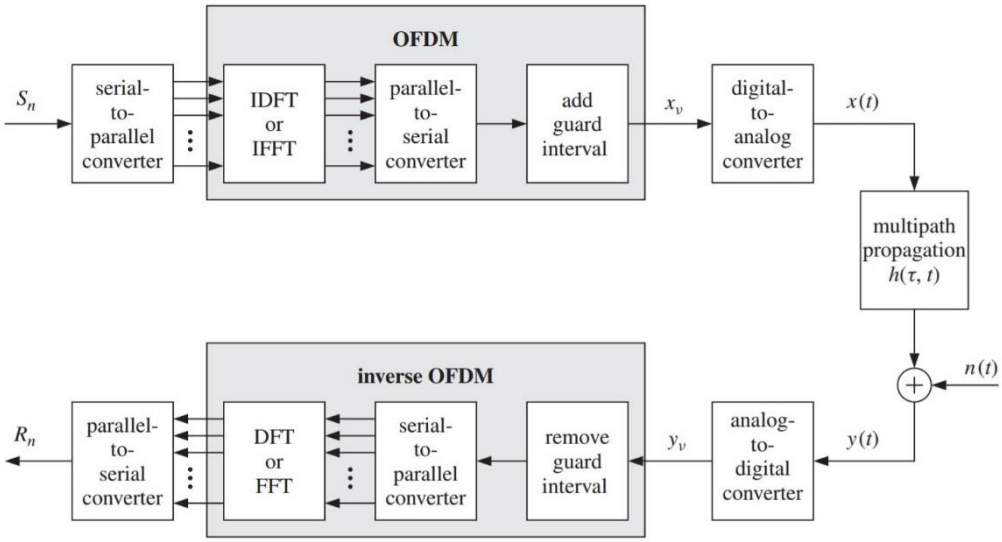

Figure 3.3 - Digital multicarrier transmission applying OFDM [33]

3.2 Basics about Signal Quantization

Quantization can be uniform or non-uniform. In both cases the quantized signal, x_Q, can be defined as the sum of the signal we want to quantize, x, and the interference error introduced, ε_Q, where the original signal has a power of P_x.

$$x = x_Q + \varepsilon_Q \,. \tag{3.5}$$

In the case of uniform quantization, L levels are separated by an uniform step Δ. If the original signal that we want to quantize has a power of P_x and a mean value equal to zero, taking continuous values between $-A_M$ and A_M, where A_M is the maximum level of the original signal, then, the step is given by:

$$\Delta = 2\frac{A_M}{L} \,. \tag{3.6}$$

The signal is equally spaced in L intervals, $[a_i, a_{i+1}[$ where $L = 2^b$ and b is the number of bits.

$$a_i = A_M\left(\frac{2i-2}{2^b} - 1\right),\ i = 1,\ldots,2^b \,. \tag{3.7}$$

The values of the signal x in each of the previous intervals are quantized to the following levels:

$$q_i = a_i + \frac{\left(a_{i+1} - a_i\right)}{2} = A_M \left(\frac{2i-1}{2^b} - 1 \right),\ i = 1, \ldots, 2^b . \tag{3.8}$$

Thus, the characteristic function is given by:

$$g_x = \begin{cases} A_M & , x \geq A_M \\ A_M \left(\frac{2i-1}{2^b} - 1 \right) & , x \in \left[a_i, a_{i+1} \right[,\ i \in \left\{ 1, \ldots, 2^b \right\} \\ -A_M & x \leq -A_M \end{cases} . \tag{3.9}$$

Considering the quantized channel as:

$$y = g\left(x\right) = \alpha x + d . \tag{3.10}$$

The average power of the useful signal is given by:

$$S = \left|\alpha\right|^2 \sigma_x^2 . \tag{3.11}$$

The average power of the quantized signal is:

$$P_{out} = E\left[\left|y\right|^2 \right] = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{+\infty} g^2\left(x\right) e^{-\frac{x^2}{2\sigma_x^2}} dx . \tag{3.12}$$

The average power error of quantization is then given by:

$$P_\varepsilon = P_{out} - S . \tag{3.13}$$

The factor α is given by:

$$\alpha = \frac{E\left[y^* x \right]}{E\left[\left|x\right|^2 \right]} = \frac{1}{\sqrt{2\pi\sigma^3}} \int_{-\infty}^{+\infty} x g\left(x\right) e^{-\frac{x^2}{2\sigma_x^2}} dx . \tag{3.14}$$

The generic characteristic is represented in Fig. 3.4 below. For the uniform quantization, the variance can be given by:

$$\sigma_\varepsilon^2 = \frac{\Delta^2}{12} . \tag{3.15}$$

where the Δ is the amplitude of each quantization level.

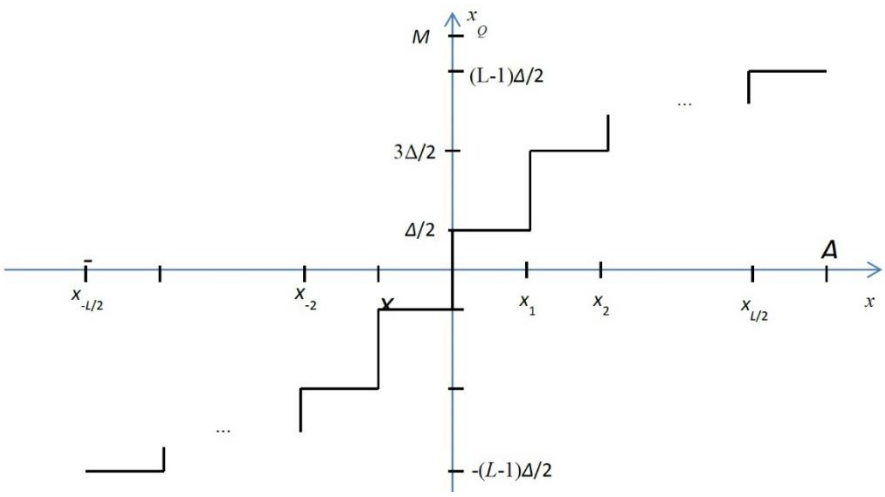

Figure 3.4 - Characteristic function for uniform quantization

For input signals whose range are infinity, quantization characteristic can be limited between $-A_M$ e A_M, where A_M is the saturation level. For this we consider ideal clipping, which corresponds to the characteristic in the Fig. 3.5 and is given by:

$$f_c(x) = \begin{cases} A_M & , x \geq A_M \\ x & , -A_M \leq x \leq A_M \\ -A_M & , x < A_M \end{cases} . \tag{3.16}$$

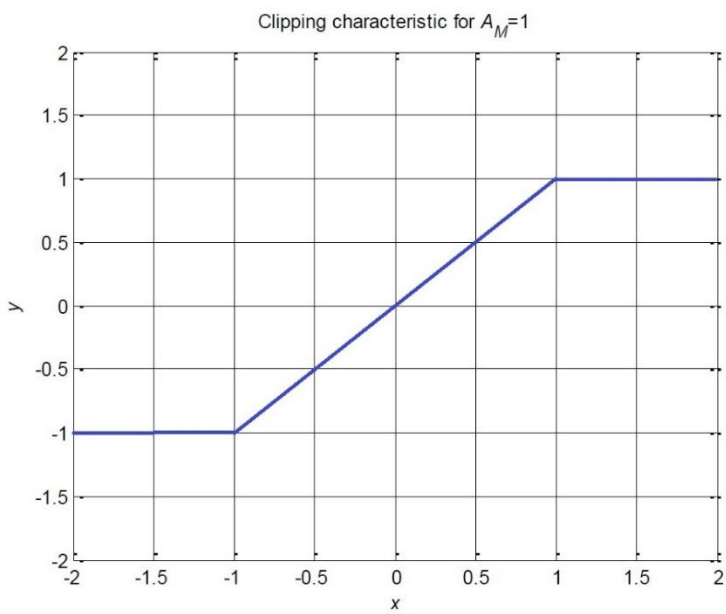

Figure 3.5 - Clipping characteristic function for saturation level of A_M

Thus, the characteristics became the ones in Fig. 3.6, for $A_M = 1$ and $b = 3$:

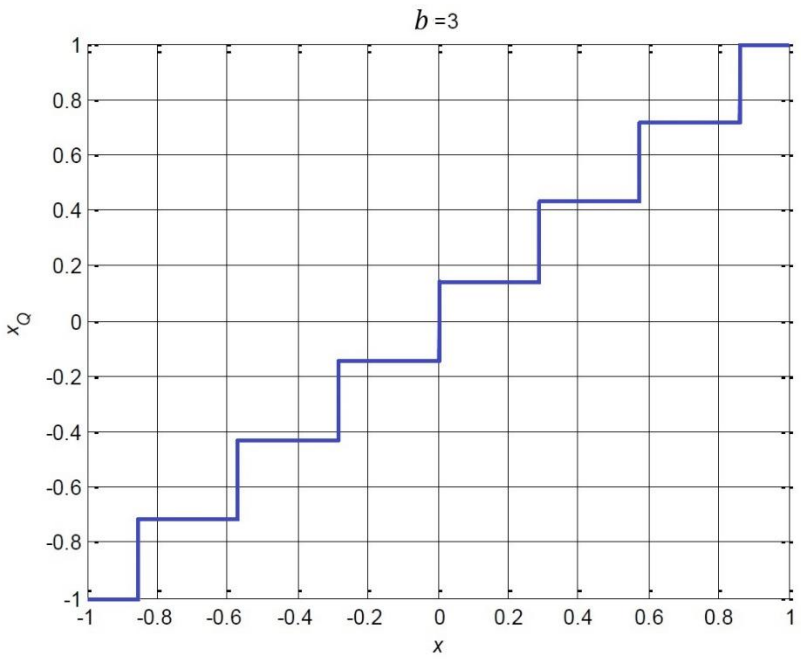

Figure 3.6 - Characteristic function for uniform quantization for a saturation level of $A_M = 1$

The signal-to-quantization-noise ratio (SQNR) is given by:

$$\text{SQNR}=\frac{P_x}{E\left[\left|x-x_Q\right|^2\right]}. \tag{3.17}$$

The SQNR depends on the saturation level as we can observe in Fig. 3.7, and it can be seen that if the number of bits b increases, the signal-to-quantization-noise ratio gets higher, so the effect of the noise on signal detection is lower.

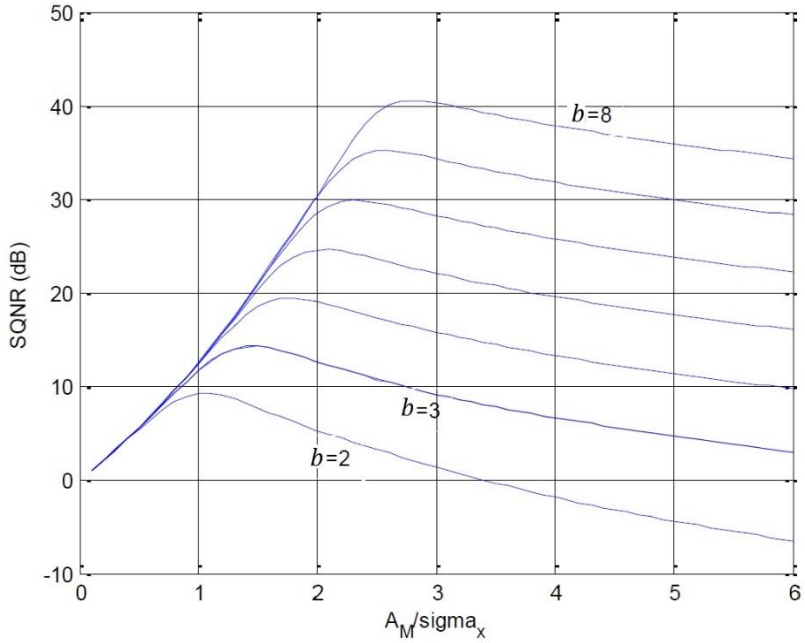

Figure 3.7 - SQNR curve in function of normalized saturation level for real uniform quantization

3.3 Random Vector Quantization

As said above, RVQ is a technique that allows a simple approach of the codebook design were the vectors are generated independently from a uniform distribution on the complex unit sphere. This approach can limit the overhead needed to feedback CSIT. However, the complexity attached to this technique's codebook can be very large as the number of transmitters and receivers increases.

In this section, we briefly describe RVQ feedback quantization technique often considered for MIMO based systems. The constructed codebook for channel direction information (CDI), defined as the normalized CSI, i.e.:

$$\boldsymbol{h}_{k,l}^{d} = \frac{\boldsymbol{h}_{k,l}}{\left|\boldsymbol{h}_{k,l}\right|} . \tag{3.18}$$

Is formed by 2^b vectors independent and identical distributed o the M -dimensional unit sphere, where b represents the number of feedback bits per OFDM symbol and user:

$$\{c_b\}, b = 1, ..., 2^b . \tag{3.19}$$

Each user quantizes its CDI to a codework in a given codebook $\boldsymbol{C}_k \in \mathbb{C}^{2xb}$ and the codebook is predetermined and known at both the BS and user sides. Partial CSI is acquired at the transmitter via finite rate feedback channel from each of the receivers.

Furthermore, we use the minimum Euclidean distance to choose the codeword closest to each channel:

$$f_{k,l} = \arg \min_{i=1,\dots,2^b} \left\| \boldsymbol{h}_{k,l}^{d} - \boldsymbol{c}_i \right\|^2 . \tag{3.20}$$

where $k = 1,\dots,K$, $l = 1,\dots,N-1$. Thus, after each user terminal having sent the index of the codeword to the BS, the BS obtain the CSI through the corresponding codebooks and using the indexes given by:

$$\boldsymbol{h}_{k,l}^{Q} = \boldsymbol{c}_{f_{k,l}} . \tag{3.21}$$

Thus, it can be design the precoder matrices. Only the CDI is sent, dismissing the channel information with this method [52].

3.4 Uniform Quantization

RVQ technique becomes impractical as the number of antennas increases, in terms of computational complexity, as for the massive MIMO case. For these cases a limited feedback technique based on Uniform Quantization can be addressed, where only a portion of the channel frequency response (CFR) and/or the channel impulse response (CIR) is quantized. The CFR is estimated at the receiver through appropriate training sequences and/or pilots. The CIR has a duration (this duration is measured in terms of number of samples) that must be smaller than the duration of the cyclic prefix, N_{CP} , which for typical OFDM implementations is much lower than N_c. Therefore, CIR must be zero for taps higher than N_{CP} ,i.e., only the first N_{CP} samples of the Cir are non-zero. When $N_c \geq N_{CP}$ it is enough to sample the CFR at a rate $0, N_c / N_{CP}, 2\mathrm{N}_c / N_{CP}, \dots$,i.e., we only need N_{CP} equally spaced samples of the CFR to obtain it without loss of information. From the samples chosen to be quantize $\left\{ h_{k,m,l'\Delta \mathrm{f}_s}^{Q} ; l' = 0,1,\dots,N_s - 1 \right\}$ with $\Delta f_s = N_c / N_s$, we obtain $\left\{ h_{k,m,l'}^{Q} ; l' = 0,1,\dots,N_s - 1 \right\}$. We consider the separate quantization of the real and imaginary parts of each of the appropriate N_{CP} samples of $\boldsymbol{h}$, leading to:

$$h_{k,m,l'}^{Q} = f_Q\left(\mathrm{Re}\left\{ h_{k,m,l',\Delta f_s} \right\}\right) + i f_Q\left(\mathrm{Im}\left\{ h_{k,m,l',\Delta f_s} \right\}\right). \tag{3.22}$$

Where $f_Q(.)$ denotes the quantization characteristics.

3.5 UQ and RVQ in a Multi-user Communication

In order to compare both channel quantization techniques defined previously, we will consider the downlink of an OFDM based system and assume a single BS equipped with M_t antennas simultaneously transmitting K independent messages to K spatially dispersed single antenna UTs sharing the same resources as shown in Fig. 3.8, where $M_t \geq K$.

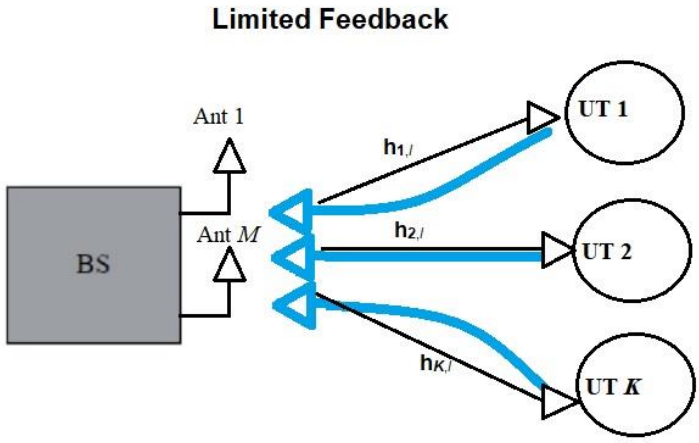

Figure 3.8 - MU - MIMO system model.

An OFDM modulation with N_c available subcarriers is employed at the BS and a linear precoding is done individually over the N_c subcarriers. The channel frequency response (CFR) associated to the link between the BS and the $k\text{th}$ user terminal for the $l\text{th}$ subcarrier is represented by the vector $\mathbf{h}_{k,l} \in \mathbf{C}^{M_t x1}$, where $k = 1,...,K$, $l = 1,...,N-1$. The quantized version of the CSI is fed back by one of the strategies addressed in the next section.

The received frequency signals are given by:

$$y_l = \mathbf{H}_l \mathbf{W}_l \mathbf{s}_l + \mathbf{n}_l . \tag{3.23}$$

where $\mathbf{H}_l = \left[\mathbf{h}_{1,l}, \mathbf{h}_{2,l}, ..., \mathbf{h}_{K,l}\right]^T$ with $\mathbf{H}_l \in \mathbf{C}^{KxM_t}$, is the equivalent channel that contains the flat Rayleigh fading coefficients. $\mathbf{s}_l$ is the data symbols vector, with $\mathbf{s}_l \in \mathbf{C}^{Kx1}$, $\mathbf{W}_l \in \mathbf{C}^{M_t xK}$ is the linear precoding matrix computed at the BS on a subcarrier l and $\mathbf{n}_l$ is the AWGN vector at the subcarrier l. The linear precoding matrix, when is taking into account the channel quantization errors, can be given by:

$$W = \alpha \left(\mathbf{H}_l^Q\right)^H \left(\mathbf{H}_l^Q \left(\mathbf{H}_l^Q\right)^H + \left(K\sigma_n^2 + \sigma_Q^2\right)\mathbf{I}_K\right)^{-1}. \tag{3.24}$$

where σ_n^2 is the variance of the quantization error and $\mathbf{I}_K = E\left[\mathbf{s}_l \mathbf{s}_l^H\right]$.

If we consider the use of b quantization bits for real and imaginary parts of each sample with UQ technique and B bits for each user of RVQ, the total number bits required for CSI quantization with UQ is $2bKM_tN_s$ bits and with RVQ is BKN_c bits [52] [53].

Fig. 3.9 presents the BER evaluation of these two addressed in this chapter in a multi-user MISO system, when using the same amount of quantization bits. The scenario has BS equipped with 2 antennas (M =2) and 2 single antennas user terminals (K =2). The main parameter used in simulations are based on LTE standards and for details see [52]. The number of quantization bits b is 4, 5 and 5 which corresponds to the values of $B = 2bM$ equals to 16, 20 and 24, respectively.

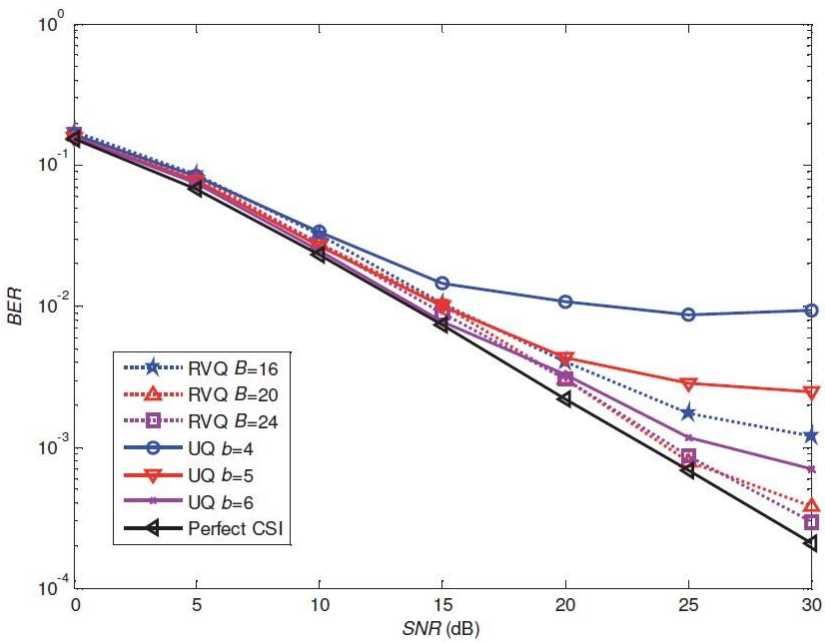

Figure 3.9 - BER performance using RVQ and UQ for $M_t = 2$ and $K = 2$ [52]

In Fig. 3.9, it can be seen that RVQ technique, despite of having higher complexity, has slightly better performances than the UQ method. As the number of quantization bits increases the difference between both techniques, UQ vs RVQ, starts to be insignificant.

Chapter 4

Millimeter Waves and Massive MIMO

The current 4G systems including LTE and WiMAX already use advanced technologies such as OFDM, MIMO, multi-user diversity, link adaptation, codes in order to achieve spectral efficiency close to theoretical limits in terms of bits per second per Hertz per cell, however, this is not sufficient for future necessities [40].

Despite all advances, it is necessary to shift to a new paradigm because of the unbridled growth of mobile telecommunications user's requirements. As we know, the frequencies spectrum in use are already in a state of overcrowding, so one essential solution in changing this so-called paradigm is the use of the millimeter waves (mmWave) frequencies.

In this third chapter, the millimeter wave ranges and their combination with the massive MIMO technologies are detailed described. The symbiotic combination of millimeter waves and massive antennas has the potential to dramatically improve the wireless access and throughput [12].

4.1 Millimeter Waves

On the current spectrum, almost all current mobile communication systems use the spectrum in the range of 300 MHz – 3GHz [11]. It is already known that this band is generally referred as *sweet spot* due to its favorable propagation conditions in commercial wireless application. However, like said before, the increasing popularity of the smart phones and other data devices, mobile data traffic is experiencing a new unparalleled growth. In order to meet this exponential growth, it is necessary to improve the air interface capacity and relocate the new spectrum.

The sub-3 GHz frequencies spectrum is becoming increasingly very crowded, but, in other hand, a vast amount of unused spectrum rage in the 3 GHz – 300 GHz has been largely unexploited for commercial wireless applications. This untouched band joint the Super-High Frequency (SHF) and Extremely-High Frequency (EHF). These two types of frequencies in this band share similar propagation characteristics and can be referred to as millimeter waves (1 – 100 mm) band. Millimeter waves spectrum would allow service providers to significantly expand the channel bandwidths far beyond the present 20 MHz channels used by 4G customers [54]. Fig. 4.1 shows the mmWave spectrum.

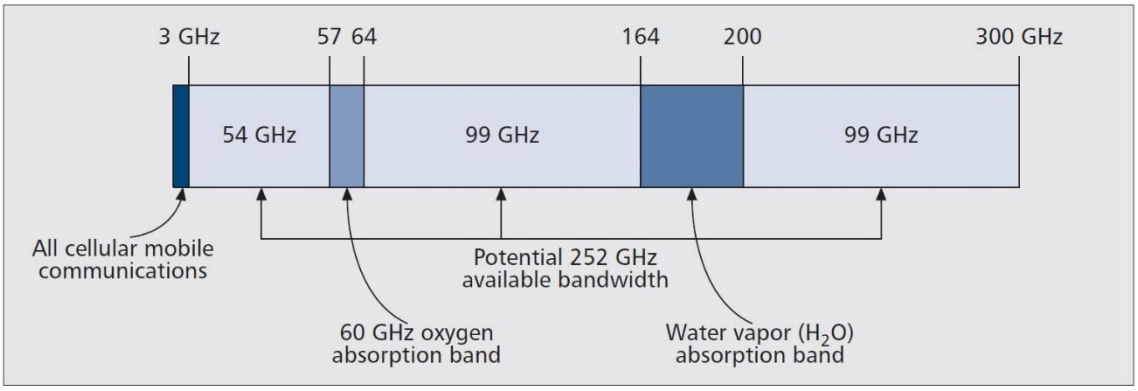

Figure 4.1 - Millimeter Wave Spectrum [11]

Just more recently there has been some interest in exploring this spectrum for short-range and fixed wireless communications. For example, the range 3.1 GHz – 10.6 GHz has proposed to enable the high data rate connectivity in personal area networks. Other example, is the oxygen absorption that has been also promoted to provide multigigabit data rates for short-range connectivity and wireless local area network. In the oxygen absorption band, it can be experienced attenuation of about 15 dB/Km. In these frequencies, the oxygen molecule absorbs electromagnetic energy. Other band with higher attenuation is the water absorption band, 164 GHz – 200 GHz. These two bands are excluded for mobile broadband applications because the transmission range will be limited by the factors above said. Additionally, the Federal Communications Commission (FCC) auctioned two licenses to the Local Multipoint Distribution Service (LMDS) start to use the frequencies from 28 GHz to 30 GHz, a service of point-to-multipoint technology in the last mile. Another bands announced by FCC were the 71 GHz - 76 GHz, the 81 GHz – 86 GHz and the 92 GHz – 95 GHz. They are referred as E-band and has become available to ultra-high speed data communication including point-to-point wireless local area networks, mobile backhaul and broadband internet access. E-band permit systems in these bands to be engineered in close proximity to one another without causing interference [11].

In terms of propagation characteristics, a general misconception among wireless developers is that free-space propagation loss depends on frequency: higher frequencies propagate less well than the lower frequencies. This misconception is the underlying assumption often used that the path loss is calculated at a specific frequency between two antennas, whose effective aperture areas increases with the wavelengths (decreases with carrier frequency). So, an antenna with larger aperture has a greater gain than a smaller one. However, if we think ahead, with the shorter wavelengths more antennas can be packed in the same area, so, for the same aperture area, in the term of free space loss, these shorter wavelengths should not have any inherent disadvantage compared to the longer wavelengths. To further, larger number of antennas enable transmitter and receiver beamforming with higher gains [11] [55].

In the mmWave, the attenuation and losses are less than a few dBs per kilometer [56] excluding, like it was already seen, in the oxygen and water absorption bands. Although, reflections and diffractions, which depends on the material, reduces the range of mmWave but facilitates the non-line of sight (NLOS) communications. Signals at lower frequencies can penetrate more easily though buildings than higher frequencies, so, it is necessary other means to have indoor coverage, like femtocells or WI-FI solutions. Another agent of propagation is the multipath. With narrow transmitter and receiver beams, the multipath components of mmWave - that helps the NLOS transmissions – are limited. Fig. 4.2 shows the attenuation for different materials.

		Attenuation (dB)		
Material	**Thickness (cm)**	**< 3 GHz [6, 8]**	**40 GHz [7]**	**60 GHz [6]**
Drywall	2.5	5.4	–	6.0
Office whiteboard	1.9	0.5	–	9.6
Clear glass	0.3/0.4	6.4	2.5	3.6
Mesh glass	0.3	7.7	–	10.2
Chipwood	1.6	–	.6	–
Wood	0.7	5.4	3.5	–
Plasterboard	1.5	–	2.9	–
Mortar	10	–	160	–
Brick wall	10	–	t178	–
Concrete	10	17.7	175	–

Figure 4.2 - Attenuation for different materials [11]

There are another factors that may cause losses like foliage losses that for mmWave are significant and, in some cases, can be a limiting impairment for propagations. The rain is another significant issue. Raindrops have similar size as the mmWave wavelengths, so, in the presence of rain, depending on the volume, a mmWave transmission can experience significant attenuation. We need to approach an alternative

communication over the cellular system when mmWave communications are been disturbed by heavy rains. However, if we consider the fact that today's cell sizes are about 200m in urban environments, it becomes clear that rain attenuation will cause only 7 dB/Km for communications at 28 GHz, which translates to only 1.4 dB of attenuation over 200m. Many researchers have confirmed that small distances (less than 1 Km), rain attenuation will present a minimal effect on the propagation of mmWave at 28 GHz to 38 GHz for small cells [57].

With the introduction of these new frequencies in the telecommunications scenario, it is necessary to ensure the mobility is guaranteed through the millimeter mobile broadband (MMB). By assuming the mobility of the receiver, we will experience the Doppler phenomena. The Doppler shift values of the incoming waves on different angles at the receiver results on the Doppler spread. In the MMB, the narrow beams at the transmitter and receiver sides will reduce angular spreads of incoming waves, what reduces the Doppler spread. Therefore, the time domain in a MMB channel is likely to be much less than in an omnidirectional antenna in a rich scattering environment [11].

Another important discussion is how to deploy the infrastructures that support the MMB. An MMB network consists of a multiple MMB BSs that cover a geographic area. To create a seamless coverage, millimeter mobile broadband base stations (MMB BSs) need to be deployed with higher density than the current macro cellular systems. In these systems, the transmission and reception are based in narrow beams, which suppress the interference between MMB BSs and can extend the range of the beam. This is opposite of what happens in the regular cellular systems and allows a significant overlap of coverage among neighboring BS's. So, the MMB BSs form a grid with a large number of nodes. The cost to connect every MMB BSs via wired network can be significant, so, to reduce the cost, we can allow some of them to connect to the backhaul via other MMB BSs. Because of the large beamforming gains, the MMB inter-BS backhaul link can be deployed in the same frequency as the MMB access link without creating much interference. This allows to achieve a higher density coverage. However, to coverage some holes and some losses, the system information, control information and the feedback can be transmitted in the previous 4G systems, thus leaving mmWave spectrum available for higher data rate communications. This is referred as hybrid MMB + 4G systems.

A key enabling technology of MMB is the beamforming that is a signal processing technique used for a directional signal transmission or reception. The direction of the signal is achieved by using adaptive transmit/receive beam patterns. When transmitting, a beamformer controls the amplitude and the phase at each transmitter antenna to create a pattern of constructive or destructive interference in the wavefront. When receiving, signals from many antennas are combined in a way that expected pattern of radiation is observed. For MMB transceivers, the small size and separation of the mmWave antennas allow a large number of antennas and thus achieve high beamforming gain in a small area. Beamforming can be implemented digital baseband, analog baseband or RF front-end, depending on the system we want to implement [11].

Given this significant jump in bandwidth and new capabilities offered by mmWave, the links between BSs and devices, as well as backhaul links between BSs, will be able to handle much greater capacity than today's 4G networks in highly populated areas [9].

4.2 Massive MIMO

Increasing the capacity and reliability of wireless communications systems through the use of multiple antennas has been an active area of research for over 20 years. MIMO wireless systems are now part of current standards and are deployed throughout the world [12]. With the evolution of smart terminals, their applications and wireless communications traffic, the capacity of wireless communications networks has to be increased in order to guarantee the Quality of Service (QoS). Therefore, it is necessary to introduce new technologies to support the demands of the exponential traffic for the next generation's wireless communications [58].

MIMO technology has attracted attention in wireless communications because can offer increases in the data throughput and link range without any additional increase in bandwidth or transmit power. So, nowadays, MIMO is already a key 4G technology and is capable of improving both bandwidth gains, multiplexing gains ad diversity gains. In order to increasing these gains, the massive MIMO concept appeared [59]. Like already said in previous chapters, this concept use hundred or even thousands of antennas, but is only feasible when using mmWave. In case of using sub 6 GHz with massive MIMO, the antennas would have to be distributed over large areas.

Bandwidth efficiency is still an important issue and has to be figured in the next technologies, but, is also the power consumptions in wireless networks that begin to worry. Carbon emissions and operator's expenditures increase year by year, so energy efficiency has become another significant issue [60] [61]. With massive MIMO technique, bandwidth efficiency can be improved since it can achieve large multiplexing gains when serving tens of user's terminals at the same time. Energy efficiency can be increased too due to the fact that the use of more antennas helps focus energy with an extremely narrow beam on small regions where the user terminals are located [62]. An example that massive MIMO systems are a key advantage in energy efficiency gains is that a user in is kind of system, with ideal CSI, can theoretically achieve the same uplink throughput as with a single-antenna BS using only 1/ N_t the required transmit power [63]. A similar scaling law can be applied in the downlink. These gains in energy efficiency can be used to help overcome the larger path losses exhibited at mmWave frequencies and extend the system's operational range.

More explicitly, massive MIMO technique refers to the system that uses hundreds of antennas to simultaneously serve dozens of users equipment that, both theoretical

and measurement results, indicate that is capable of significantly improving the bandwidth efficiency and simultaneously reducing the transmit power [64]. Because of all that, massive MIMO is seen as a candidate technique for 5G.

Traditional MIMO technology can only adjust signal transmission in the horizontal dimension. In order to exploit the vertical dimension of signal propagation, antennas arrays (AA), such rectangular, spherical and cylindrical arrays were studied. MIMO with these arrays can adjust azimuth and elevation angles and can propagate signals in three-dimensional (3D) space. This is referred to 3D MIMMO. Massive MIMO adopts rectangular, spherical or cylindrical AAs, therefore, 3D MIMO with massive number of antennas can be seen as practical deployment means of massive MIMO.

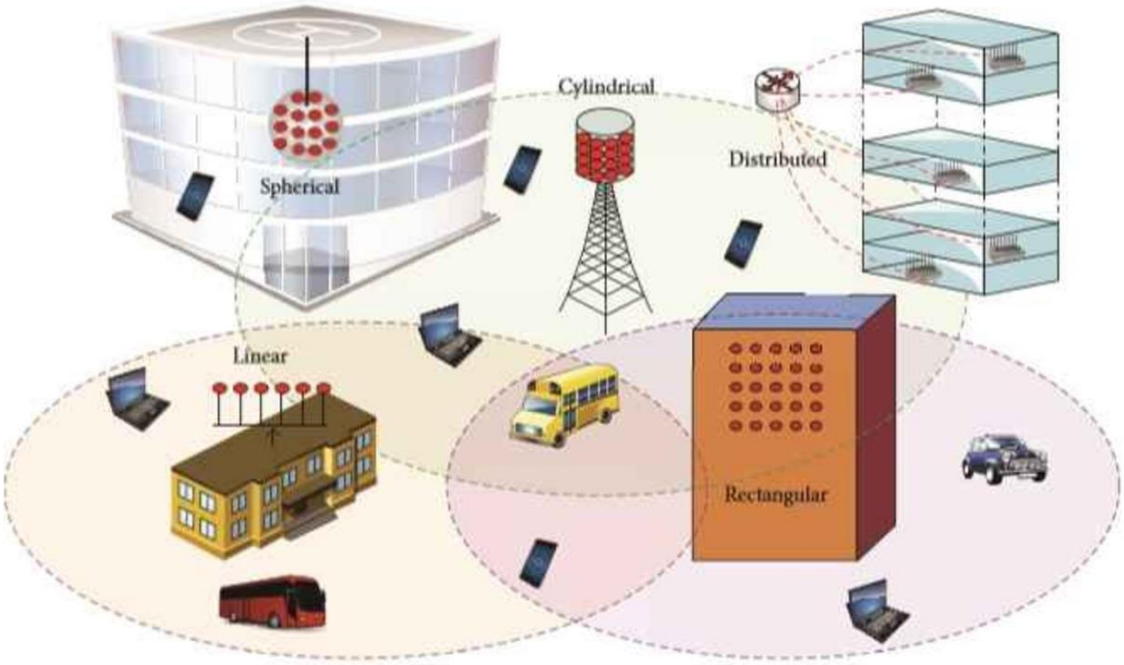

Figure 4.3 - Various antennas configurations

Fig. 4.3 illustrates the various antennas configurations. The linear AA is an example of a 2D AAs. The spherical, cylindrical and rectangular belongs to the 3D AAs. The last ones are more realistic for practical systems. The distributed AA is mainly used either inside buildings or for outdoor cooperation. The linear AA is mostly assumed just in theoretical analysis of realistic measurements [58]. An aesthetically pleasing method is to deploy massive MIMO as a part of building's facade [65].

The antenna configuration directly affects the characteristics of a massive MIMO channel. The linear AA gives rise both to non-stationary channel characteristics and to near-filed effects, while spherical, cylindrical and rectangular AAs are capable of accurately directing the beam propagation ins the 3D space. Correlation Based

Stochastic Model (CBSM) is the type of model that has been used for evaluating the performance of wireless communications systems due to its simplicity.

There are two typical scenarios with massive MIMO: Homogeneous (HomoNet) and Heterogeneous networks (HetNet). The first one, HomoNet, only use macro cells and, on the contrary, the second ones, HetNet, is deployed woth both macro and small cells. In the Fig. 4.4, in a very simply manner, it is shown the features of each type.

	Type	Description	Feature
HomoNet	Case 1A	Multi-layer sectorization	Easy to implement High multiplexing gains
	Case 1B	Adaptive beamforming	Narrow beam and little interference Flexible with elevation and azimuth
	Case 1C	Large-scale cooperation	Coverage enhancement Efficient cooperation
HetNet	Case 2A	Wireless backhaul	Low infrastructure costs Flexible and extendible
	Case 2B	Hotspot coverage	High throughput Superior elevation resolution
	Case 2C	Dynamic cell	Adjusting cell adaptively Balancing network load

Figure 4.4 – Features of typical applications scenarios [58]

One important issue is the inter-cell interference coordination (ICIC). Cellular communication systems suffer from inter-cell interference at the cell boundaries, especially when all the channels are fully reused in adjacent cells. Interference mitigation techniques are needed to alleviating this kind of interference. As we approach massive MIMO as one of the key technologies to support 5G, it is known that a large scale of antenna arrays can provide additional spatial DoF. Therefore, ICIC for massive MIMO systems is able to exploit the spatial DoF for mitigating interference by nulling certain spatial direction to neighboring cell [66].

In the HomoNet context, as for the ICIC, the beamforming with massive MIMO systems, can help eliminate the interference because massive MIMO has the capability to dynamically adapt the shape of the vertical forming pattern to the user terminal at different locations. This result in an increased system throughput.

The HetNet is an attractive mean of increasing achievable network capacity and enhancing the coverage area. In this type of network, small cells, as a tier, are capable of providing hotspot capacity enhancements, while macro cells, as another tier, are responsible for large area coverage in support of high mobility user terminals. However,

they may interfere with each other if they use the same time-frequency resources without carefully coordination. Fortunately, like said above, large scale antenna arrays can provide an additional spatial DoF for multiplexing the data of several users on the same time-frequency resource. Furthermore, it can concentrate the radiation energy precisely on the intended user terminal, reducing both intra and inter-tier interference. So, massive MIMO systems are capable of supporting cooperation in an implicit way between different layers in the HetNet for the sake of improving the overall system performance.

4.3 Massive MIMO with Millimeter Waves

In popular frequency band below 6 GHz, the physical size of large scale antenna arrays is excessive to be feasibly installed in a user terminal. In order to avoid thus problem, massive MIMO systems are likely to operate in the millimeter wave band [12]. However, the channel measurements indicate propagation losses that are severe in the mmWave band, hence the propagation distance becomes very short, resulting in small cells. To obtain data rates and satisfactory coverage, the ultra-dense deployment of small cells operating in the mmWave region is one of the solutions envisaged for next generation wireless communications networks.

4.3.1 Hybrid Structures

There are many challenging issues regarding the implementation of digital beamforming in massive MIMO systems: complexity, energy consumption and cost. Full digital beamforming can yield the optimal performance. However, when a large number of antennas are deployed, implementing the same number of transceivers may not be feasible due to excessive demand on real-time signal. A beamforming structure with much lower number of digital transceivers than the number of antennas will therefore be more practical and cost effective to deploy. One way to reduce the number of transceivers is via analog beamforming structure, where each transceiver is connected with multiple active antennas, and the signal phase on each antenna is controlled via a network of analog phase shifters. Normally, each transceiver generates one beam toward one user in analog beamforming. When the user number simultaneously served are smaller than the antenna number -what happens often in massive MIMO systems -, the transceiver number can be designed to be much smaller than the antenna number. However, if the users are not adequately spatially separated, there may be severe inter-user interference. So, digital beamforming over transceivers can then be utilized to achieve multiple data stream precoding on top of analog beamforming to further enhance the performance.

Like it is seen in Fig. 4.5, there are two hybrid structures that has been under scrutiny, with N being the transceiver number and $NM = M_t$ being the number of antennas.

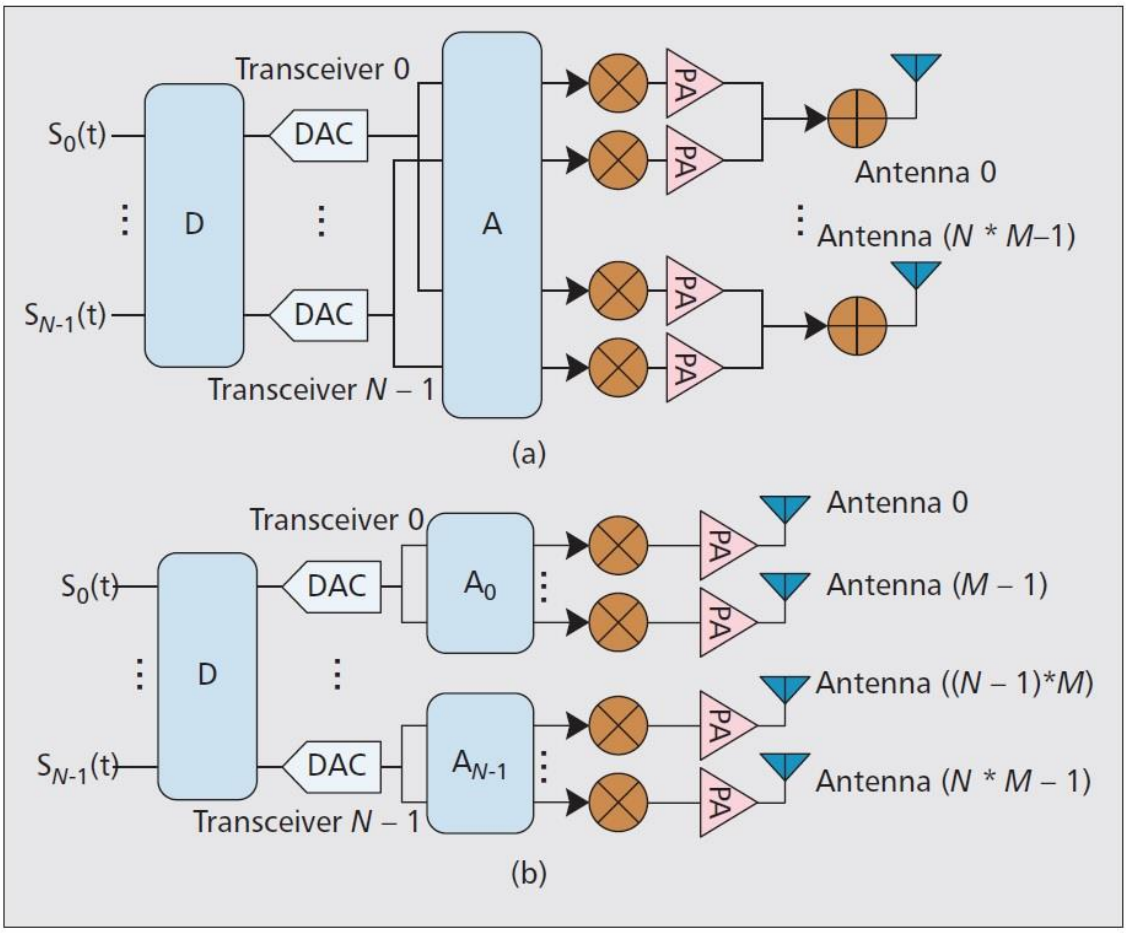

Figure 4.5 - Hybrid beamforming structures a) full connected b) N by M hybrid beamforming structure [67]

In the structure illustrated in Fig. 4.5 a), the N transceivers are connected with all antennas. So, the transmitted signal on each digital transceiver goes through NM RF paths and summed up before being connected with each antenna element. Analog beamforming is performed over NM RF paths per transceiver and digital beamforming can then be performed over N transceivers. This structure can achieve full gain over transceivers. However, the complexity of this structure is very high, e.g., the total number of RF paths is N^2M .

Illustrated in Fig. 4.5 b) is an NM hybrid beamforming structure where each of the N transceivers are connected to M antennas. Analog beamforming is performed over only M RF paths in each transceiver and digital beamforming is performed over

N transceivers. This structure is more practical for BS antenna deployments in the current cellular systems, like 3G and 4G systems, where each transceiver is generally connected to a column of antennas. With active antennas on each RF path, elevation beamforming can be performed by applying different phases to each antenna in each column. Compared to 4.5 a) structure, the beamforming gain per transceiver is $1/N$ the gain of 4.5 a), however with much reduced complexity with the total number of RF paths being NM.

There are many issues yet to be addressed regarding the hybrid beamforming structures, including reference signals designs, digital/analog beamforming designs, etc. The difficulty in the design of these structures is mainly due to the fact that is difficult to have CSI knowledge at BS transmitter side. The mapping between transceivers and antennas makes channel estimation in both downlink and uplink more complicated [67] [68].

4.3.2 mmWave Channel Model

Taking into account the system model presented in [69], we consider the single user mmWave system with M_t transmitter antennas and M_r receiver antennas using the clustered channel model with N_{cl} scattering clusters, each of which contribute N_{ray} propagation paths to the channel matrix $\mathbf{H}$. Therefore, the discrete-time narrowband channel $\mathbf{H}$ can be written as

$$\mathbf{H} = \gamma \sum_{i,1} \alpha_{il} \Lambda_r \left(\phi_{il}^r, \theta_{il}^r\right) \Lambda_t \left(\phi_{il}^t, \theta_{il}^t\right) \mathrm{a}_r \left(\phi_{il}^r, \theta_{il}^r\right) \mathrm{a}_t \left(\phi_{il}^t, \theta_{il}^t\right)^* . \tag{4.1}$$

where γ is a normalization factor:

$$\gamma = \sqrt{\frac{M_t M_r}{N_{cl} N_{ray}}} . \tag{4.2}$$

α_{il} is the complex gain of the lth ray in the ith scattering cluster, whereas $\phi_{il}^r\left(\theta_{il}^r\right)$ and $\phi_{il}^t\left(\theta_{il}^t\right)$ are its azimuth (elevation) angles of arrival and departure respectively. The functions $\Lambda_r\left(\phi_{il}^r, \theta_{il}^r\right)$ and $\Lambda_t\left(\phi_{il}^t, \theta_{il}^t\right)$ represent the receive and transmit antenna element gain at the corresponding angles of arrival and departure, respectively. Finally, the vectors $\mathrm{a}_r\left(\phi_{il}^r, \theta_{il}^r\right)$ and $\mathrm{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)$ represent the normalized receive and transmit array response vectors at an azimuth (elevation) angle of $\phi_{il}^r\left(\theta_{il}^r\right)$ and $\phi_{il}^t\left(\theta_{il}^t\right)$ respectively. We assume that α_{il} are independent and identical distributed $CN = \left(0, \sigma_{\alpha,i}^2\right)$

where $\sigma_{\alpha,i}^2$ represents the average power of the ith cluster. The average cluster powers are such that:

$$\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \gamma \,. \tag{4.3}$$

Where γ is a normalization that satisfies $\mathrm{E}\left[\left|\mathbf{H}\right|^2\right] = M_t M_r$. The N_{ray} azimuth and elevation angles of departure, ϕ_{il}^t and θ_{il}^t, within the cluster i are assumed to be randomly distributed with a uniformly random mean cluster angle of ϕ_i^t and θ_i^t respectively and a constant angular spread (standard deviation) of σ_{ϕ^t} and σ_{θ^t} respectively. The azimuth and elevation angles of arrival, ϕ_{il}^r and θ_{il}^r, are again randomly distributed with mean cluster angles of $\left(\phi_i^r, \theta_i^r\right)$ and angular spreads $\left(\sigma_{\phi^r}, \sigma_{\theta^r}\right)$. The angles of arrival and departure in clustered channel models act by the Laplacian distribution since has been found to be a good fit for a variety of propagation scenarios.

For the functions $\Lambda_t\left(\phi_{il}^t, \theta_{il}^t\right)$ and $\Lambda_r\left(\phi_{il}^r, \theta_{il}^r\right)$, a number of parametrized mathematical models have been proposed, for example if the transmitter's antenna elements are modeled as being ideal sectored elements, $\Lambda_t\left(\phi_{il}^t, \theta_{il}^t\right)$ would be given by:

$$\Lambda_t\left(\phi_{il}^t, \theta_{il}^t\right) = \begin{cases} 1 & \forall \phi_{il}^t \in \left[\phi_{\min}^t, \phi_{\max}^t\right], \forall \theta_{il}^t \in \left[\theta_{\min}^t, \theta_{\max}^t\right] \\ 0 & \text{otherwise} \end{cases} . \tag{4.4}$$

where we have assumed unit gain over the sector defined by $\phi_l^t \in \left[\phi_{\min}^t, \phi_{\max}^t\right]$ and $\theta_l^t \in \left[\theta_{\min}^t, \theta_{\max}^t\right]$ without loss of generality. The receive antenna element gain is defined similarly.

The array response vectors $\mathrm{a}_r\left(\phi_{il}^r, \theta_{il}^r\right)$ and $\mathrm{a}_t\left(\phi_{il}^t, \theta_{il}^t\right)$ are a function of the receiver and the transmitter antenna array structure only and, subsequently, independent of the antenna element properties. There are two examples of commonly used antenna arrays for completeness. For an N-element uniform linear array (ULA) on the y-axis, the array response vector can be written as:

$$\mathrm{a}_{\mathrm{ULAy}}\left(\phi\right) = \frac{1}{\sqrt{N}}\left[1, e^{jkd\sin(\phi)}, \ldots, e^{j(N-1)kd\sin(\phi)}\right]^T . \tag{4.5}$$

where $k = \dfrac{2\pi}{\lambda}$ and d is the inter-element spacing. Note that θ is not included in the arguments of the $\mathrm{a}_{\mathrm{ULAy}}$ as the array's response is invariant in the elevation domain. In

the case of a uniform planar array (UPA) in the yz-plane with N_y and N_z elements on the y and z axes respectively, the array response vector is given by:

$$\mathrm{a}_{\mathrm{UPAy}}(\phi,\theta)=\frac{1}{\sqrt{N}}\left[1,\ldots,e^{jkd\left(m\sin(\phi)\sin(\theta)+n\cos(\theta)\right)},\ldots,e^{jkd\left(\left(N_y-1\right)\sin(\phi)\sin(\theta)+\left(N_z-1\right)\cos(\theta)\right)}\right]^T. \quad (4.6)$$

where $0\le m\le N_y$ and $0\le n\le N_z$ are the y and z indices of an antenna element respectively and the antenna array size is $N=N_yN_z$.

UPA is used in mmWave beamforming since they yield smaller antenna dimensions, facilitate packing more antenna elements in a reasonably sized array and can enable beamforming in the elevation domain (3D beamforming).

Chapter 5

Limited Feedback Strategy for Massive MIMO mmWave Systems

New and efficient spatial processing techniques can be developed in systems that combine mmWave and massive MIMO, such as beamforming/precoding and spatial multiplexing at the transmitter and/or receiver sides.

In the last chapters, we discussed that one of challenges in the deployment of mmWave in wireless applications is the difficulty in having the CSIT due to the large amount of data required for feedback. The knowledge of CSI is crucial to improve the downlink performance, and assuming that the CSIT is perfect is not realistic in many practical scenarios. This problem is even more significant in massive MIMO based systems since the terminals are equipped with a large number of antennas and therefore a huge amount of channels need to fed back from the receiver to the transmitter, increasing the overall system signaling [70].

We have seen that the mmWave communications with massive MIMO has some hardware restrictions, as the high cost and power consumption of some mmWave mixed signal components that makes it difficult to have fully dedicated RF chain for each antenna as in conventional MIMO systems. To overcome these limitations, hybrid analog/digital architecture solutions have been discussed, where some signal processing is done at the digital level and some left to the analog domain [67], [71]. They achieve a balance between the only analog solution (with low complexity and limited performance) and the only digital option (with high complexity but better performance).

In this chapter, it is proposed and evaluated an efficient limited feedback strategy for a hybrid mmWave massive MIMO OFDM system, where only a part of the parameters associated to the complex link channel are quantized and fed back. We consider a transmitter employing a hybrid analog-digital precoding/beamforming and a receiver equipped with a hybrid analog-digital equalizer/combiner. The limited feedback strategy addressed consists of quantization in time domain of only some channel parameters, such as amplitudes, angles of arrival (AoA), angles of departure (AoD), and feedback them to the transmitter, thus requiring a lower limited feedback overhead than with the case of full channel quantization. Then, based on these parameters, the space-frequency channel is reconstructed and used to compute the analog-digital precoders. Finally, the system is evaluated under the proposed CSI quantization strategy and the results are compared with the case where perfect channel is known. This will give us insights about the impact of imperfect CSIT on future hybrid mmWave massive MIMO systems. Moreover, a hybrid analog-digital equalizer is designed based on the analog-beamforming algorithm proposed in [72].

5.1 Hybrid mmWave massive MIMO platform

The mmWave massive MIMO system model considered in this work is presented in the section below, followed by the channel characterization and transmitter and receiver design definition.

5.1.1 System Model

The mmWave massive MIMO system model considered for a cellular OFDM-based system is shown in Fig. 5.1 and Fig. 5.2. We consider a hybrid analog/digital based architecture for a single user with M_t transmitter antennas and M_r receiver antennas, with a digital baseband precoder followed by a constraint RF precoding, where both transmitter and receiver are equipped with a large antenna array and the number of RF chains is lower than the number of antennas. An OFDM modulation with N_c available subcarriers is employed and a cyclic prefix duration N_{CP} is considered, where N_{CP} is usually much lower than N_c .

The received signal at transmitter on subcarrier k is given by:

$$\boldsymbol{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k . \tag{5.1}$$

where $k = 1 \ldots N_c$; $\boldsymbol{y}_k \in \mathbb{C}^{M_r x1}$ denotes the received signal vector, $\mathbf{x}_k \in \mathbb{C}^{M_t x1}$ is the transmitted signal, $\mathbf{n}_k \in \mathbb{C}^{M_r x1}$ is a zero mean Gaussian noise with variance σ_n^2 and $\mathbf{H}_k \in \mathbb{C}^{M_r x M_t}$ is the channel matrix for subcarrier k .

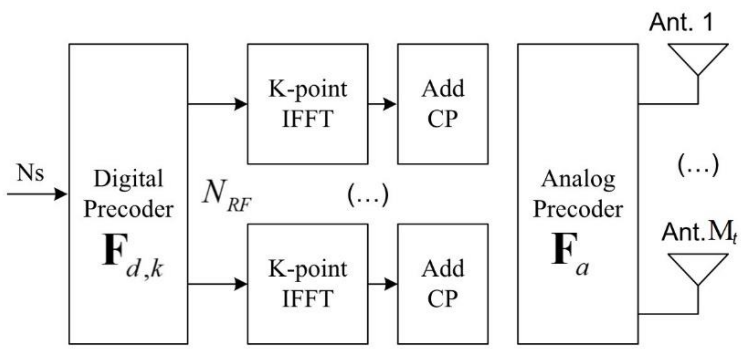

Figure 5.1 - Transmitter Block Diagram

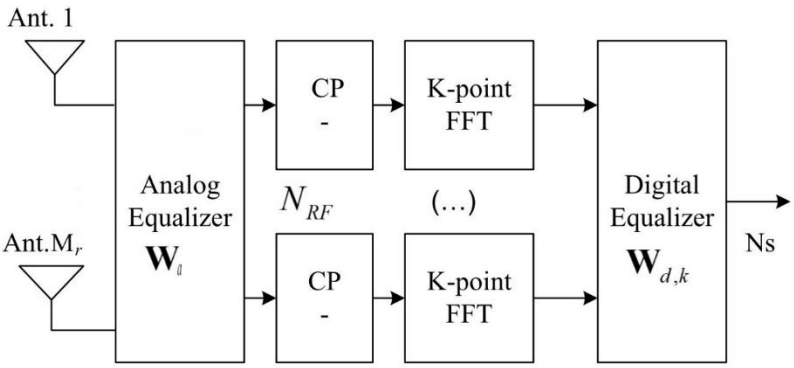

Figure 5.2 - Receiver Block Diagram

5.1.2 Channel Model

In order to accurately model a broadband mmWave channel, that combines tightly packed antenna arrays, the channel model considered follows the clustered frequency-selective mmWave channel model discussed in [72], is given by the sum of the contribution of N_{cl} clusters, each of which contributing with N_{ray} propagation paths to the channel matrix with a cyclic prefix duration of N_{CP}. The channel at subcarrier k is expressed as [72]:

$$\mathbf{H}_k = \sum_{d=0}^{N_{CP}-1} \mathbf{H}[d] e^{-k\frac{2\pi k}{N_{cl}}d} . \tag{5.2}$$

where $\mathbf{H}[d]$ is the delay-d MIMO matrix, with D taps that can be written as:

$$\mathbf{H}[d] = \gamma \sum_{i=1}^{N_{cl}} \sum_{l=1}^{M_r} \alpha_{i,l} p\left(dT - \tau_{i,l}\right) \mathbf{A}_r(i,l) \mathbf{A}_t(i,l). \tag{5.3}$$

$\alpha_{i,l}$ and $\tau_{i,l}$ correspond to the paths gains and the relative time delay of the lth ray in the ith scattering cluster, respectively. γ is a normalization factor such that

$$\gamma = \sqrt{\frac{M_t M_r}{N_{ray} N_{cl}}} . \tag{5.4}$$

$p(\tau)$ is a pulse-shaping function for T-spaced signaling evaluated at τ seconds. $\mathbf{A}_t$ and $\mathbf{A}_r$ are the matrices of array response vectors at the transmitter and receiver, respectively formed by:

$$\mathbf{A}_t = \left[\mathbf{a}_t\left(\theta_{t1,1}\right),...,\mathbf{a}_t\left(\theta_{tN_{cl},\mathrm{N}_{ray}}\right)\right]. \tag{5.5}$$

and

$$\mathbf{A}_r = \left[\mathbf{a}_r\left(\theta_{r1,1}\right),...,\mathbf{a}_r\left(\theta_{rN_{cl},\mathrm{N}_{ray}}\right)\right]. \tag{5.6}$$

whereas $\theta_{i,l}^r$ and $\theta_{i,l}^t$ are the azimuth angles of arrival and departure. The array response vectors for an N_w-element uniform linear array is given by:

$$\mathbf{a}_w\left(\theta_{wi,l}\right) = \frac{1}{\sqrt{N_w}}\left[1, e^{j\frac{2\pi d_a}{\lambda}\sin\left(\theta_{wi,l}\right)},...,e^{j\frac{2\pi\left(N_w-1\right)d_a}{\lambda}\sin\left(\theta_{wi,l}\right)}\right]^T. \tag{5.7}$$

with $w \in \{r,t\}$ and d_a is the inter-element spacing.

Moreover, we assume that the duration of the delay-d MIMO channel is smaller than the cyclic prefix duration N_{CP}. Therefore, we have $\mathbf{H}[d]=0$ for $d > N_{CP}$.

5.1.3 Transmitter Design

The transmitter processing is decomposed into two parts, the digital baseband and the analog circuitry that are modeled mathematically by precoder matrices $\mathbf{F}_a \in \mathbb{C}^{M_t x N_{RF}}$ and $\mathbf{F}_{d,k} \in \mathbb{C}^{N_{RF} x N_s}$, respectively. The digital part has N_{RF} transmit chains, with $N_s \le N_{RF} \le M_t$, where N_s is the length-K data symbols blocks. Due to hardware constricts, the analog part is implemented using a matrix of analog phase shifters, which force all elements of matrix $\mathbf{F}_a$ to have equal norm $\left(\left|\mathbf{F}_a\left(i,l\right)\right|^2 = M_t^{-1}\right)$. The transmit signal is given by:

$$\mathbf{x}_k = \mathbf{F}_a\mathbf{F}_{d,k}\mathbf{s}_k\ . \tag{5.8}$$

where $\mathbf{s}_k \in \mathbb{C}^{N_s x 1}$ is the data vector for subcarrier $k = 1,...,N_c$. The hybrid analog-digital precoder is based on Gram-Schmidt orthogonalization, that leads to a near-optimal complexity proposed in [72]. Note that the analog precoder is constant over all the subcarriers while the digital precoder is computed on a per subcarrier basis. The pseudo code of this algorithm is presented in Table 5.1.

Table 5.1 – Hybrid analog-digital precoding Algorithm

Algorithm 1. Hybrid Precoder design

$$\left[\sim, \mathbf{S}_k, \mathbf{V}_k\right] = svd\left(\mathbf{H}_k\right)$$

$$\mathbf{S} = \left[\mathbf{S}_1\left(1:N_{RF}, 1:N_{RF}\right), ..., \mathbf{S}_{N_c}\left(1:N_{RF}, 1:N_{RF}\right)\right]$$

$$\mathbf{V} = \left[\mathbf{V}_1\left(:, 1:N_{RF}\right), ..., \mathbf{V}_{N_c}\left(:, 1:N_{RF}\right)\right]$$

$$\prod = \mathbf{V}\mathbf{S}^H$$

$$\mathbf{M} = \mathbf{I}_{M_t}$$

$$\mathbf{F}_a = EmptyMatrix$$

for $r = 1, ..., N_{RF}$

$$\boldsymbol{\psi} = \prod^H \mathbf{M}\mathbf{A}_t$$

$$i_{opt} = \arg \max_{i=1,...,N_{ray}N_{cl}} \left|\boldsymbol{\psi}^H \boldsymbol{\psi}\right|_{i,i}$$

$$\mathbf{F}_a = \left[\mathbf{F}_a\left(\mathbf{A}_t\left(:, i_{opt}\right)\right)\right]$$

$$\mathbf{M} = \mathbf{I}_{M_t} - \mathbf{F}_a^H\left(\mathbf{F}_a^H\mathbf{F}_a\right)^{-1}\mathbf{F}_a^H$$

end

$$\mathbf{H}_{RF,k} = \mathbf{H}_k\mathbf{F}_a$$

$$\mathbf{F}_{d,k} = \mathbf{H}_{RF,k}^H\left(\mathbf{H}_{RF,k}\mathbf{H}_{RF,k}^H\right)^{-1}$$

$$\mathbf{F}_{d,k} = \frac{\mathbf{F}_{d,k}}{\left\|\mathbf{F}_{d,k}\right\|_F}\sqrt{N_{RF}}$$

5.1.4 Receiver Design

At the receiver, it was designed a hybrid analog-digital equalizer. The design is based on a similar principle used to design the precoders. It should be mentioned that in [72] only precoders were designed since the aim was to evaluate the mutual information. In this chapter, the mmWave massive MIMO system is evaluated in terms of BER and therefore a hybrid equalizer should be designed. The received signal is firstly processed through the analog phase shifters, modeled by matrix $\mathbf{W}_a \in \mathbb{C}^{N_{RF} x M_r}$, then follows the baseband processing composed of N_{RF} processing chains. All elements of the matrix

$\mathbf{W}_a$ must have equal norm $\left(\left|\mathbf{W}_a\left(j,l\right)\right| = M_r^{-1}\right)$. The signal first passes through a linear filter $\mathbf{W}_{d,k} \in \mathbb{C}^{N_s x N_{RF}}$. The underlying filtered received signal is the given by $\mathbf{W}_{d,k}\mathbf{W}_a\mathbf{y}_k$. The proposed analog and digital matrices are described in Table 5.2. As for hybrid precoder, the analog part of the equalizer is constant over all the subcarriers while the digital part is computed for each subcarrier. The matrix $\mathbf{W}_{ad,k}$ represents the full digital MMSE equalizer matrix used for the case where a RF chain per antenna is assumed.

Table 5.2 –Equalizer Algorithm designed for hybrid analog-digital architecture

<u>Algorithm 2. Proposed Equalizer design</u>

$$\mathbf{H}_{eq,k} = \mathbf{H}_k\mathbf{F}_a\mathbf{F}_{d,k}$$

$$\mathbf{R}_{y,k} = \mathbf{H}_{q,k}\mathbf{H}_{eq,k}^H + \sigma_n^2\sigma_s^{-2}\mathbf{I}_{M_t}$$

$$\mathbf{W}_{ad,k} = \left(\mathbf{H}_{eq,k}\mathbf{H}_{eq,k}^H + \sigma_n^2\sigma_s^{-2}\mathbf{I}_{M_t}\right)\mathbf{H}_{eq,k}^H$$

$$\mathbf{W}_{ad} = \sum_{k=1}^{N_c}\mathbf{W}_{ad,k}$$

$$\mathbf{M} = \mathbf{I}_{M_r}$$

$$\mathbf{W}_a = EmptyMatrix$$

for $r = 1,\ldots,N_{RF}$

$$\boldsymbol{\psi} = \mathbf{W}_{ad}\mathbf{M}\mathbf{A}_t$$

$$i_{opt} = \arg\max_{i=1,\ldots,N_{ray}N_{cl}}\left|\boldsymbol{\psi}^H\boldsymbol{\psi}\right|_{i,i}$$

$$\mathbf{W}_a = \left[\mathbf{W}_a\left(\mathbf{A}_t\left(:,i_{opt}\right)\right)^H\right]$$

$$\mathbf{M} = \mathbf{I}_{M_r} - \mathbf{W}_a^H\left(\mathbf{W}_a\mathbf{W}_a^H\right)^{-1}\mathbf{W}_a$$

end

$$\mathbf{W}_{d,k} = \mathbf{W}_{ad}\mathbf{R}_{y_k}\mathbf{W}_a^H\left(\mathbf{W}_a\mathbf{R}_{y_k}\mathbf{W}_a^H\right)^{-1}$$

5.2 Millimeter Wave Channel Quantization

The CSI is needed at both transmitter and receiver sides to compute the hybrid analog-digital precoders and equalizers/combiners. For the sake of simplicity, we assume perfect channel estimation at the receiver side. In practice, the channel can be estimated at the receiver through appropriate training sequences and/or pilots [73], however, these methods for CSI estimation are out of this work. From the previous section, we can see that to compute the analog-digital precoders, we need to know matrix $\mathbf{H}_k$ for all subcarriers and the transmit array response $\mathbf{A}_t$. More specifically, to compute the digital precoder we need matrix $\mathbf{H}_k$ and to compute the analog precoder we need matrix $\mathbf{A}_t$ and thus, this information should be fed back from the receiver to the transmitter. This would imply a quantization and feedback of $M_t M_r + N_{ray} N_{cl}$ complex values per subcarrier, which would result in $2N_c b_f \left(M_t M_r + N_{ray} N_{cl} \right)$ bits of overhead, where b_f is the number of quantization bits per real or imaginary part of channel component.

In order to reduce the feedback overhead it is proposed to quantize and fed back just some parameters that are characteristic of the channel, followed by channel reconstruction at the transmitter and precoder computation. Notice that from (5.3), it can be seen that the channel matrix $\mathbf{H}_k$ depends on three main parts: the transmit array response, $\mathbf{A}_t$, the Rayleigh fading complex coefficients and the receive array response, $\mathbf{A}_r$. Therefore, we quantize some parameters individually, i.e., we propose to quantize uniformly the complex path gains $\alpha_{i,l}$, and the real values of the azimuth angles of arrival and departure $\theta_{ri,l}$ and $\theta_{ti,l,k}$ respectively, with $i = 1,...,N_{cl}$ and $l = 1,...,N_{ray}$, for each of the taps of delay-d MIMO channel, obtaining the quantized variables:

$$\alpha_{i,l} = f_{Q\alpha}\left(\mathrm{Re}\left\{\alpha_{i,l}\right\}\right) + i f_{Q\alpha}\left(\mathrm{Im}\left\{\alpha_{i,l}\right\}\right). \tag{5.9}$$

$$\theta_{Qri,l} = f_{Qr}\left(\theta_{ri,l}\right). \tag{5.10}$$

$$\theta_{Qti,l} = f_{Qt}\left(\theta_{ti,l}\right). \tag{5.11}$$

f_{Qr} and f_{Qt} are the quantization characteristics functions based on uniform quantization, using b_α, b_r and b_t bits, respectively, and where the levels are equally spaced in $\left[-\alpha_c, \alpha_c\right]$, $\left[\theta_{ri} - \theta_{rc}, \theta_{ri} + \theta_{rc}\right]$ and $\left[\theta_{ti} - \theta_{tc}, \theta_{ti} - \theta_{tc}\right]$, respectively with α_c, θ_{tc} and θ_{rc} being the clipping values for amplitude and azimuth angles of arrival and departure, respectively, and θ_{ti} and θ_{ri} are the mean AoD ad AoA for cluster i. For the azimuth

angles of arrival and departure, the clipping values are given by one-half of the receiver and transmitter sector angles.

After quantization, these parameters are fed back to the transmitter, where they are then used to reconstruct the matrices of array response vectors of transmitter and receiver, defines as

$$\mathbf{A}_{tQ} = \left[\mathbf{a}_t \left(\theta_{Qt1,1} \right), ..., \mathbf{a}_t \left(\theta_{QtN_{cl},N_{ray}} \right) \right]. \tag{5.12}$$

$$\mathbf{A}_{rQ} = \left[\mathbf{a}_r \left(\theta_{Qr1,1} \right), ..., \mathbf{a}_r \left(\theta_{QrN_{cl},N_{ray}} \right) \right]. \tag{5.13}$$

respectively and the channel matrix reconstructed using these quantized variables in (5.3).

To obtain the precoder matrix, we follow the steps described in Table 5.1, Algorithm 1, using $\mathbf{H}_Q$ and $\mathbf{A}_{tQ}$. The quantization overhead for the proposed strategy is then given by $N_{CP} N_{ray} N_{cl} \left(2b_\alpha + b_r + b_t \right)$. In a particular scenario that has $N_{cl} = 6$ clusters, each with $N_{ray} = 5$ rays; the number of antennas at transmitter and receiver are $M_t = 16$ and $M_r = 4$, respectively; the number of subcarriers is set to $N_c = 128$ and the cyclic prefix has a duration of $N_{CP} = 32$; if we consider the number of quantization bits equal for all parameter, $b_f = b_\alpha = b_r = b_t = 8$, we obtain the comparison between the two quantization strategies illustrated in Table 5.3.

Table 5.3 - Comparison between two quantization strategies

Full quantization strategy	$2N_c b_f \left(M_t M_r + N_{ray} N_{cl} \right) = 192512$ bits
Proposed quantization strategy	$N_{CP} N_{ray} N_{cl} \left(2b_\alpha + b_r + b_t \right) = 30720$ bits

Note that the quantization of channel parameters will affect the combining process in the receiver, being important to conclude about how the quantization errors will affect the overall system performance. A block diagram showing these dependencies is presented in the Fig. 5.3.

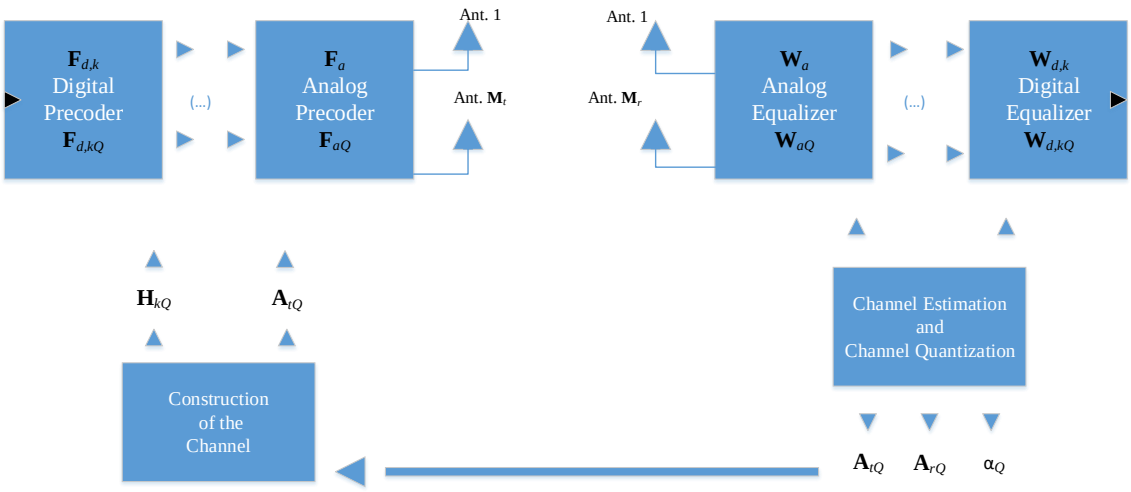

Figure 5.3 - Channel Quantization Block Diagram

5.3 Performance Results

In this section, we present a set of performance results for the limited feedback method suggested for mmWave model described above. Our scenario has $N_{cl} = 6$ clusters, each with $N_{ray} = 5$ rays, with Laplacian distributed azimuth angles of arrival and departure. The antenna inter-element spacing d_a is assumed to be half-wavelength. The number of antennas at transmitter and receiver are $M_t = 16$ and $M_r = 4$, respectively, and the number of RF chains is $N_{RF} = 4$.The number of subcarriers is set to $N_c = 128$ and the cyclic prefix has a duration of $N_{CP} = 32$. The number of data symbols transmitted per subcarrier is set to $N_s = 4$.

We assume that $\alpha_{i,l}$ are independent and identically distributed $CN(0, \sigma_{\alpha,i}^2)$, where $\sigma_{\alpha,i}^2$ defines the average power of the ith cluster. The average power of all N_{cl} clusters is the same and such that $\sum_{i=1}^{N_{cl}} \sigma_{\alpha,i}^2 = \gamma$, where γ is the normalization factor in (5.4). The N_{ray} azimuth angles of departure and arrival within cluster i, $\theta_{ti,l}$ and $\theta_{ri,l}$, respectively, are assumed to follow a Laplacian distribution, with a uniformly-random mean cluster angle of θ_{ti} and θ_{ri}, and a constant angular spread $\sigma = 10^{\circ}$.

The performance metric considered is the BER, which is presented as a function of the E_b / N_0, with E_b denoting the average bit energy and N_0 denoting the one-sided noise power spectral density.

In Fig. 5.4, we observe the performance of the hybrid mmWave system when we just quantize the transmitter phases with a different number of bits and for the same clipping value ($\sigma_{tc} = 3\sigma$). As expected the performance increases with overhead rate, taking values close to the perfect CSIT with just 8 bits per subcarrier.

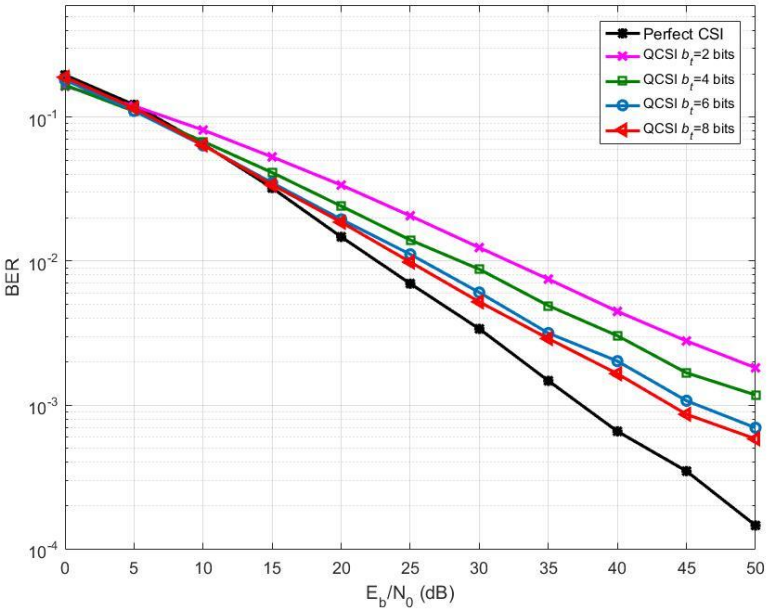

Figure 5.4 - Performance of the proposed hybrid mmWave system for quantization of AoD with and different number of quantization bits

In order to observe the performance degradation with the clipping value, Fig. 5.5 and 5.6 and 5.7 present different values for these parameter, when the number of quantization bits is 6, 8 and 10, respectively.

We can observe in Fig. 5.5 that when we quantize the transmitter phases with 6 bits, the difference between using $\sigma_{tc} = 3\sigma$ or $\sigma_{tc} = 5\sigma$ in the performance of the hybrid mmWave is almost negligible. However, if the quantization bits for this parameter increase for 8 and 10 bits - Fig.5.6 and Fig. 5.7, respectively - the performance gets better by using $\sigma_{tc} = 5\sigma$.

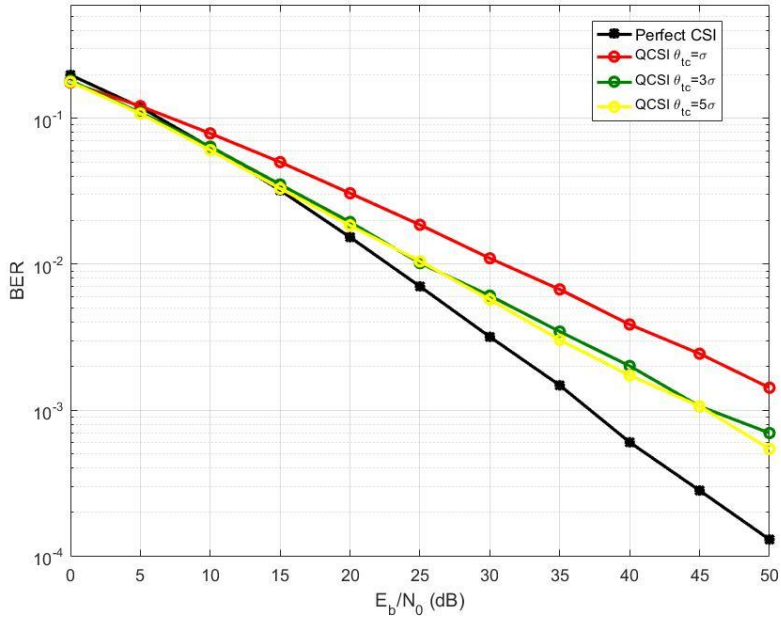

Figure 5.5 - Performance of the proposed hybrid mmWave system for quantization of transmitter phases with $b_t = 6$ and different values for clipping

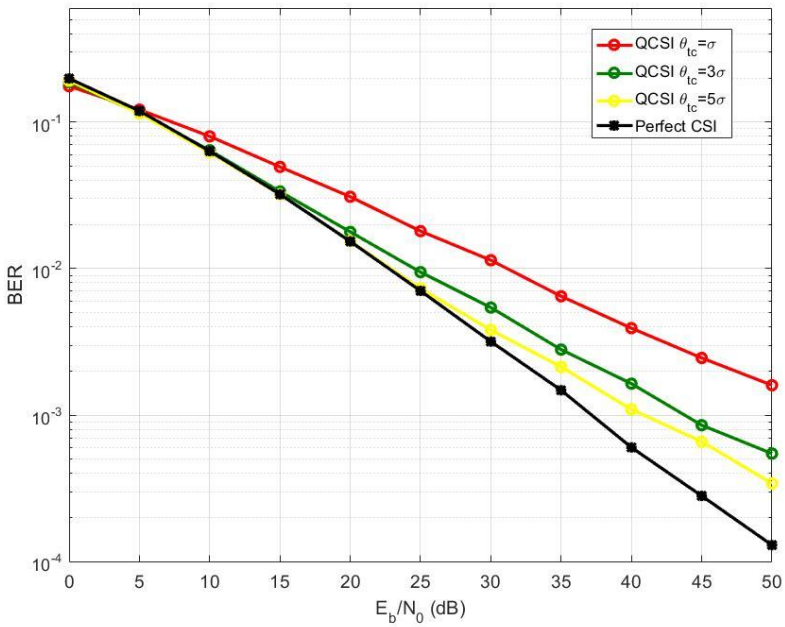

Figure 5.6 - Performance of the proposed hybrid mmWave system for quantization of transmitter phases with $b_t = 8$ and different values for clipping

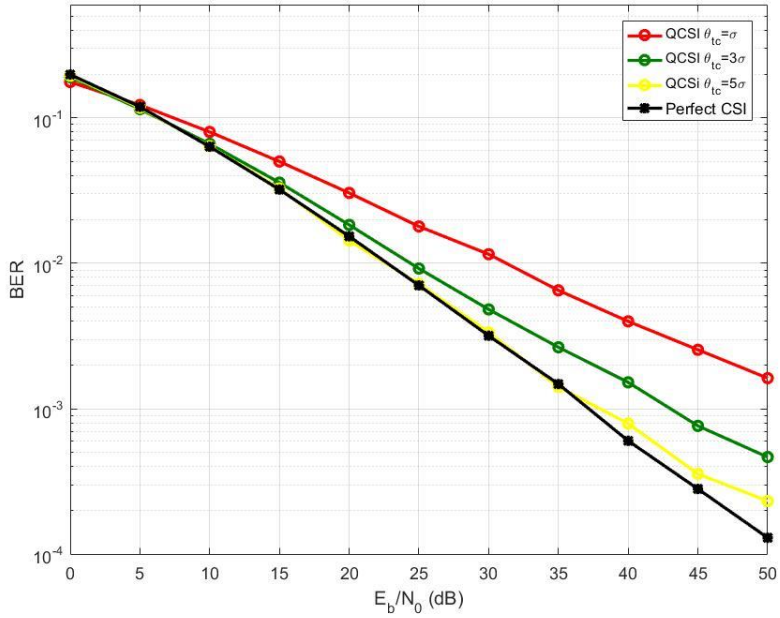

Figure 5.7 - Performance of the proposed hybrid mmWave system for quantization of transmitter phases with $b_t = 10$ and different values for clipping

In the Fig. 5.8, the curves refer to just channel amplitude values quantization when using 2 for the clipping value and having different number of quantization bits. We observe that with 8 quantization bits, the limited feedback strategy method has a performance close to the perfect CSI one.

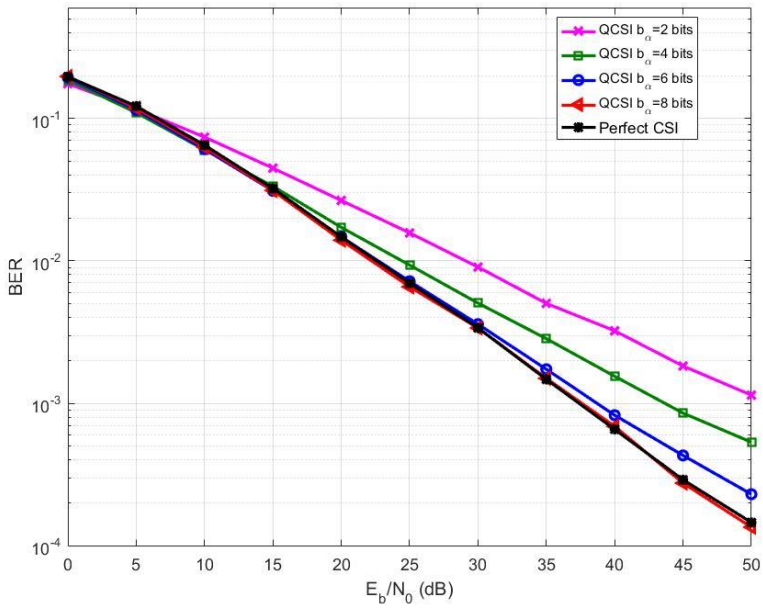

Figure 5.8 - Performance of the proposed hybrid mmWave system for quantization of channel amplitude with $\alpha_c = 2$

In Fig. 5.9, the case where three channel parameters are quantized is considered: transmitter and receiver phases, and channel amplitude. The number of quantization bits for AoA and AoD is 10 bits and the clipping is set in 5σ, while the complex amplitudes were quantized with 6 quantization bits and $\alpha_c = 2$. We define these parameters values after analyzing the graphs in Fig. 5.5, Fig. 5.6 and Fig. 5.7. The performance of the scheme with the proposed limited feedback strategy is close to the one with perfect CSIT, with just 2 dB of difference for a target BER of 10^{-3}. The channel overhead in this case would be of about 13% of the overhead used if the full channel were quantized and fed back in the frequency domain, for the same conditions.

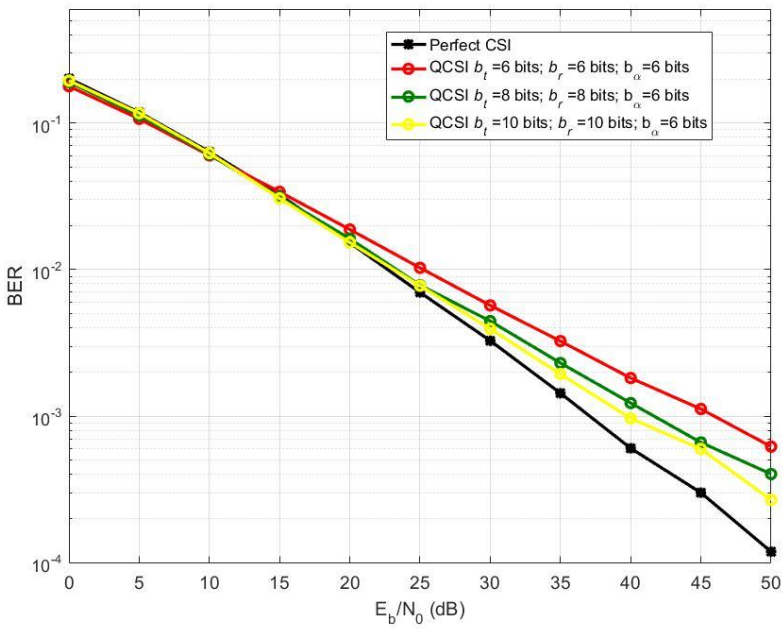

Figure 5.9 - Performance of the proposed hybrid mmWave system for quantization of transmitter and receiver phases with both 6,8 and 10 quantization bits and amplitude with 6 quantization bits, and $\theta_{tc} = 5\sigma$, $\theta_{rc} = 5\sigma$ and $\alpha_c = 2$

Chapter 6

Conclusions and Future Work

Nowadays, the current mobile device users require new demands like higher throughput and reliable communications and that made the efficiency in current telecommunications systems near limit. It is needed a new paradigm shift and this is where 5G comes in, to make difference. The two enabling key technologies are the use of the mmWave combined with massive MIMO, allowing both transmitter and receiver to have a large array of antennas. With this comes new challenges as the need to know of CSIT. In this dissertation, we have seen that assuming perfect CSIT is not realistic in practical scenarios, thus a limited feedback strategy for a hybrid mmWave massive MIMO OFDM system is proposed, where only part of the parameters associated to the complex link channel are quantized and fed back.

6.1 Conclusions

This dissertation began by presenting a summary of the evolution of cellular communications systems, from the 1st Generation to the 4th Generation. Then, a briefly overview toward to the forthcoming communication system, the 5th Generation, 5G. In Chapter 1 it was seen that through evolution of telecommunication technologies, mobile devices had become an aspect of our daily live that no one can already live without. We can almost say that nowadays, mobile devices are a crucial part for our human social life.

In Chapter 2, we have discussed how multiple antenna systems can achieve more throughput and coverage. Techniques like diversity at both sides and spatial multiplexing were approached.

In Chapter 3, OFDM modulation schemes are approached. We also introduced the quantization techniques, like Random Vector Quantization and Uniform Quantization.

In Chapter 4, firstly we introduced the mmWave and massive MIMO technologies and then the combination of the two key technologies of the 5G systems, mmWave and massive MIMO. In this chapter, we also described the hybrid architectures where the number of antennas is lower than the number of RF chains and the signal processing is done in two different domains, analog and digital. This allows the use of a massive number of antennas to perform highly directional beams and consequently maximize the received signal energy and overcome the path loss in mmWave.

In the Chapter 5, we proposed a limited feedback strategy designed for a single user OFDM mmWave massive MIMO system, where the transmitter and the receiver employs a hybrid analog-digital beamforming and equalizer. The limited feedback method is based on uniform quantization of some channel parameters that are fed back and used to compute the hybrid precoder. The proposed techniques were evaluated under practical scenarios described in Chapter 5.

The overall conclusion is, as expected, that the imperfect CSI has a strong impact on hybrid massive MIMO mmWave based systems. However, the proposed quantization strategy is quite efficient, achieving a performance close to the one obtained with perfect CSIT. The feedback overhead reduction is achieved without the need to increase the complexity of the system.

In one of the practical scenarios, we just quantized the transmitter phases with a different number of bits for the same clipping value. The values were close to the perfect CSIT with just 8 bits for each imaginary and real component of each element for the quantization of AoD parameter.

The performance degradation with the clipping values were also observed for both AoD and channel amplitude, by varying the number of quantization bits. As expected as higher is the clipping value, for the same number of bits, the system performance gets better, presenting a non-linear behavior in this variation. These simulations were useful to choose the optimal clipping values depending on the quantization values used for each of the parameters that would be quantized.

When transmitter and receiver phases and channel amplitude are quantized, the performance of the scheme with the proposed quantization strategy approached here is close to the one with perfect CSIT. With only a few bits of quantization – 6, 8 and 10 quantization bits for transmitter and receiver phases and 6 for amplitude per real or imaginary component of each quantized element -, performances get close to the perfect CSI curve, while achieving large reductions for the channel overhead.

6.2 Future Work

Regarding the future work, in upcoming researches it would be interesting to investigate in following points:

- Consider a multi user scenario and compare the quantization performance with single user scenario used in this work
- Compare the quantization technique used in this work with other quantization techniques as the RVQ with previous compression.
- Compare the performance between the hybrid analog-digital structures used in this work with other architectures, e.g. the partial connected one discussed in Chapter 4.

Bibliography

[1] E. Dahlman, S. Parkvall, J. Sköld and P. Beming, 3G Evolution: HSPA and LTE for Mobile Broadband, Elsevier Ltd., First Edition, 2007.

[2] A. Jian, A. Rapjut and S. Dixit, "Evolution of Wireless Communication," *Internation Journal of Engineering Technology Science and Research,* vol. 2, no. 4, pp. 88-90, April 2015.

[3] P. Sharma, "Evolution of Mobile Wireless Communication Networks-1G to 5G as well as Future Prospective of Next Generation Communication Network," *International Journal of Computer Science and Mobile Computing ,* vol. 2, no. 8, pp. 47-53, August 2013.

[4] R. S. a. A. Garg, "Digital Society from 1G to 5G: A comparative Study," *International Journal of Application of Innovation in Engineering & Management ,* vol. 3, no. 2, pp. 186-193, February 2014.

[5] M. A. U. Gawas, "An Overview on Evolution of Mobile Wireless Communication Networks: 1G-6G," *International Journal on Recent and Innovation Trends in Computing and Communication,* vol. 3, no. 5, pp. 3130-3133, May 2015.

[6] "Distribution of global mobile subscribers from 2008 to 2020, by technology (in percent)," Statista, Complete Source Details, 2017. [Online]. Available: https://www.statista.com/statistics/227122/distribution-of-mobile-subscribers-worldwide-by-technology/. [Accessed 1 August 2017].

[7] E. Dahlman, S. Parkvall and J. Skold, 4G: LTE/LTE-Advanced for Mobile Broadband, Elsevier Ltd., Second Edition, 2014.

[8] C. Cox, An Introduction to LTE: LTE, LTE-Advanced, SAE and 4G Mobile Communications, Wiley, March 2012.

[9] T. S. Rappaport, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," *IEEE Access (Volume: 1),* pp. 335-349, 10 May 2013.

[10] T. P. M. P. a. D. M. M. Martin Whitehead, "European Mobile Industry Observatory 2011," GSMA London Office, London, November 2011.

[11] F. Khan and Z. Pi, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine,* vol. 49, no. 6, June 2011.

[12] A. L. Swindlehurst, E. Ayanoglu, P. Heydari and F. Capolino, "Millimeter-Wave Massive MIMO: The Next Wireless Revolution?," *IEEE Communications Magazine,* vol. 52, no. 9, pp. 56-62, September 2014.

[13] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong and J. Yao, "5G on the Horizon: Key Challenges for the Radio-Access Network," *IEEE Vehicular Technology Magazine,* vol. 8, no. 3, pp. 47-53, September 2013.

[14] C.-X. Wang, F. Haider, i. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine,* vol. 52, no. 2, pp. 122 - 130, February 2014.

[15] A. Aydar, A. Borg and e. al, "tutorialspoint simply easy learning," tutorialspoint, 2017. [Online]. Available: https://www.tutorialspoint.com/5g/5g_introduction.htm. [Accessed 1 August 2017].

[16] GSMA, The Mobile Economy 2017, GSM Association, 2017.

[17] T. Bai, A. Alkhateeb and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine ,* vol. 52, no. 9, pp. 70-77, September 2014.

[18] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic and R. W. Heath, "MIMO Precoding and Combining Solutions for Millimeter-Wave Systems," *IEEE Communications Magazine ,* vol. 52, no. 12, pp. 122-131, December 2014.

[19] C. K. Au-yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Transactions on Wireless Communications,* vol. 6, no. 2, February 2007.

[20] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," *European Transaction on Telecommunications,* vol. 10, no. 6, pp. 585-595, November/December 1999.

[21] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal,* vol. 1, no. 2, pp. 41-59, Autumn 1996.

[22] V. Tarokh, H. Jafarkhani and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory ,* vol. 45, no. 5, pp. 1456 - 1467, July 1999.

[23] R. Heath and A. Paulraj, "Switching between spatial multiplexing and transmit diversity based on constellation distance," *Allerton Conference Commuication, Control and Computing ,* Octuber 2000.

[24] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory ,* vol. 49, no. 5, pp. 1073-1096, May 2003.

[25] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj and H. V. Poor, MIMO Wireless Communications, Cambridge University Press, 2007.

[26] I. Poole, "MIMO Formats - SISO, SIMO, MISO, MU-MIMO," Radio-Electronics.com, [Online]. Available: http://www.radio-electronics.com/info/antennas/mimo/formats-siso-simo-miso-mimo.php. [Accessed 1 August 2017].

[27] M. Mohaisen, Y. Wang and K. Chang, *Multiple antenna technologies,* Spetember 2009.

[28] M. Jankiraman, Space-Time codes and MIMO systems, Artech Hosue Inc, 2004.

[29] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, Cambridge University Press, 2005.

[30] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.

[31] S. Sesia, I. Toufik and M. Baker, LTE - The UMTS Long Term Evolution: From Theory to Practice, Wiley, 2011 (Second Edition).

[32] A. Sezgin, *Space-Time Codes for MIMO Systems: Quasi-Orthogonal Design and Concatenation,* Elektrotechnik und Informatik der Technischen Universit¨at Berlin, 2005.

[33] K. Fazel and S. Kaiser, Multi-Carrier and Spread Spectrum Systems: From OFDM and MC-CDMA to LTE and WiMAX, Wiley, 2008 (Second Edition).

[34] A. SIlva and A. Gameiro, *Comunicação Sem Fios,* Mestrado Integrado em Electrónica e Telecomunicações, Edição 2015-2016.

[35] B. Vucetic and J. Yuan, Space-Time Coding, Wiley, 2003.

[36] Y. Huang and K. Boyle, ANTENNAS From Theory to Practice, Wiley, 2008.

[37] A. Sibille, C. Oestges and A. Zanella, MIMO: From Theory to Implementation, Elsevier Ltd, 2011.

[38] H. Jafarkhani, "A Quasi-Orthogonal Space–Time Block Code," *IEEE Transcations on Communicatios,* vol. 49, no. 1, pp. 1-4, January 2001.

[39] F. Khan, LTE for 4G Mobile Broadband: Air Interface Technologies and Performance, Cambridge University Press, 2009.

[40] D. Azevedo, *Comparação do desempenho de Arqitecturas Híbridas para comunicações na banda das ondas milimétricas,* Portugal: Universidade de Aveiro, Master's Thesis 2016.

[41] G. Anjos, *Técnicas de Processamento MIMO para 4G,* Portugal: Universidade de Aveiro, Master's Thesis 2013.

[42] A. Molisch, Wireless Communications, Wiley, 2011 (Second Edition).

[43] D. Gesbert, M. Kountouris, R. W. H. Jr and C.-B. Chae, "From Single User to Multiuser Communications: Shifting the MIMO Paradigm," *IEEE Signal Processing Magazine,* vol. 24, no. 5, pp. 36 - 46, 2007.

[44] G. Foschini and M. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications,* vol. 6, no. 3, p. 311–335, March 1998.

[45] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Transactions on Signal Processing,* vol. 50, no. 5, pp. 1051 - 1064, May 2002.

[46] H. Sampath, P. Stoica and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Transactions on Communications,* vol. 49, no. 12, pp. 2198 - 2206, December 2001.

[47] W. Santipach and M. Honig, "Asymptotic capacity of beamforming with limited feedback," in *International Symposium onInformation Theory, 2004. ISIT 2004. Proceedings.*, Chicago, IL, USA, July 2004.

[48] W. Santipach and M. Honig, "Asymptotic performance of MIMO wireless channels with limited feedback," in *IEEE Military Communications Conference, 2003. MILCOM 2003*, Boston, MA, USA, USA, May 2004.

[49] J. Choi, D. J. Love and U. Madhow, "Limited feedback in massive MIMO systems: Exploiting channel correlations via noncoherent trellis-coded quantization," in *2013 47th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA, March 2013.

[50] Srikanth, Kumaran, Manikandan and Murugesapandian, "Orthogonal Frequency Division Multiple Acess: Is it the Multiple access system the Future?," [Online]. Available: http://www.au-kbc.org/comm/Docs/Tutorils/OFDMA_BCW_cv6.pdf. [Accessed 15 8 2017].

[51] N. C. U. Department of Compute Science, "Department of Compute Science, National Chengchi University," [Online]. Available:http://www.cs.nccu.edu.tw/~jang/teaching/NextMobCom_files/Orthogonal%20Frequency%20Division%20Multiple%20Access.pdf. [Accessed 2017 August 1].

[52] S. Teodoro, A. Silva, R. Dinis and A. Gameiro, "Interpolated UQ vs RVQ as Limited Feedback for Broadband Channels," in *Proceedings of European Wireless 2015; 21th European Wireless Conference*, Budapest, Hungary, Hungary, May 2015.

[53] S. Teodoro, A. Silva, R. Dinis and A. Gameiro, "Low-Bit Rate Feedback Strategies for Iterative IA-Precoded MIMO-OFDM-Based Systems," *Volume 2014 (2014),* vol. 2014, p. 11, 11 February 2014.

[54] T. S. Rappaport, J. N. Murdock and F. Gutierrez, "State of the Art in 60-GHz Integrated Circuits and Systems for Wireless Communications," *Proceedings of the IEEE,* vol. 99, no. 8, pp. 1390 - 1436, July 2011.

[55] D. M. Pozar, Microwave Engineering, Wiley, 2005 (Third Edition).

[56] F. C. COMMISSION, "Millimeter Wave Propagation: Spectrum Management Implications," *OFFICE OF ENGINEERING AND TECHNOLOGY,* p. Bulletin Number 70, July 1997.

[57] Z. Qingling and J. Li, "Rain Attenuation in Millimeter Wave Ranges," in *2006 7th International Symposium on Antennas, Propagation & EM Theory*, Guilin, China, October 2006.

[58] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang and L. Hanzo, "Survey of Large-Scale MIMO Systems," *IEEE Communications Surveys & Tutorials,* vol. 17, no. 3, pp. 1738 - 1760, April 2015.

[59] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE Journal of Selected Topics in Signal Processing,* vol. 8, no. 5, pp. 742 - 758, April 2014.

[60] V. K. Bhargava and A. Leon-Garcia, "Green cellular networks: A survey, some research issues and challenges," in *2012 26th Biennial Symposium on Communications (QBSC)*, Kingston, ON, Canada, May 2012.

[61] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys & Tutorials,* vol. 15, no. 1, pp. 167 - 178, February 2012.

[62] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Transactions on Wireless Communications,* vol. 9, no. 11, pp. 3590 - 3600, October 2010.

[63] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Transactions on Communications,* vol. 61, no. 4, pp. 1436 - 1449, February 2013.

[64] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine,* vol. 52, no. 2, pp. 186 - 195, February 2014.

[65] HUAWEI, "Active Antenna System: Utilizing the Full Potential of Radio Sources in the Spatial Domain," Huawei Technologies Co, 2012.

[66] A. Adhikary, E. A. Safadi and G. Caire, "Massive MIMO and Inter-Tier Interference Coordination," in *ITA Workshop*, San Diego, CA, USA, February 2010.

[67] S. Han, C.-l. I, Z. Xu and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Communications Magazine,* vol. 53, no. 1, pp. 186 - 194, January 2015 .

[68] S. Han, C. L. I, C. Rowell, Z. Xu, S. Wang and Z. Pan, "Large scale antenna system with hybrid digital and analog beamforming structure," in *2014 IEEE International Conference on Communications Workshops (ICC)*, Sydney, NSW, Australia, June 2014.

[69] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi and R. W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Transactions on Wireless Communications,* vol. 13, no. 3, pp. 1499 - 1513, January 2014.

[70] S. Teodoro, A. Lopes, R. Magueta, D. Castanheira, A. Silva, R. Dinis and A. Gameiro, "Performance Evaluation of a Frequency Selective Millimeter Wave System with Limited Feedback," Universidade de Aveiro, 2017.

[71] R. Magueta, D. Castanheira, A. Silva, R. Dinis and A. Gameiro, "Hybrid Iterative Space-Time Equalization for Multi-User mmW Massive MIMO Systems," *IEEE Transactions on Communications,* vol. 65, no. 2, pp. 608 - 620, November 2016.

[72] A. Alkhateeb and R. W. Heath, "Frequency Selective Hybrid Precoding for Limited Feedback Millimeter Wave Systems," *IEEE Transactions on Communications,* vol. 64, no. 5, pp. 1801 - 1818, May 2016.

[73] D. Neves, C. Ribeiro, A. Silva and A. Gameiro, "Channel Estimation Schemes for OFDM Relay-Assisted Systems," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, Barcelona, Spain, April 2009.

[74] G. Foschini, G. Golden, R. Valenzuela and P. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE Journal on Selected Areas in Communications,* vol. 17, no. 11, pp. 1841 - 1852, November 1999.

[75] O. Tirkkonen, A. Boariu and A. Hottinen, "Minimal non-orthogonality rate 1 space-time block code for 3+ Tx antennas," in *2000 IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications. ISSTA 2000. Proceedings (Cat. No.00TH8536)*, Parsippany, NJ, USA, USA, September 2000.

[76] H. Yin, D. Gesbert, M. Filippou and Y. Liu, "A Coordinated Approach to Channel Estimation in Large-Scale Multiple-Antenna Systems," *IEEE Journal on Selected Areas in Communications,* vol. 31, no. 2, pp. 264 - 273, January 2013.

[77] A. Adhikary, E. A. Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport and A. F. Molisch, "Joint Spatial Division and Multiplexing for mm-Wave Channels," *IEEE Journal on Selected Areas in Communications,* vol. 32, no. 6, pp. 1239 - 1255, May 2014.

[78] A. Narula, M. Lopez, M. Trott and G. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications,* vol. 16, no. 8, pp. 1423 - 1436, Octuber 1998.

[79] R. Heath and A. Paulraj, "A simple scheme for transmit diversity using partial channel feedback," in *Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)*, Pacific Grove, CA, USA, November 1998.

[80] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications,* vol. 26, no. 8, October 2008.

[81] C.-B. Chae, D. Mazzarese, T. Inoue and R. W. Heath, "Coordinated Beamforming for the Multiuser MIMO Broadcast Channel With Limited Feedforward," *IEEE Transactions on Signal Processing,* vol. 56, no. 12, pp. 6044 - 6056, December 2008.