Instituto Universitário de Lisboa

Departamento de Ciências e Tecnologias da Informação

Physical Layer Wireless Techniques for Improved Network Throughput: Network Coding, MIMO, and Full-Duplex

Filipe Manuel Alves Ennes Ferreira

Dissertação submetida como requisito parcial para obtenção do grau de
Mestre em Engenharia de Telecomunicações e Informática

Orientador:

Prof. Doutor Francisco A. T. B. N. Monteiro

ISCTE-IUL

Outubro de 2014

Resumo

Esta dissertação propõe a junção de conceitos que permitem ir para lá da simples configuração de rede com *two-way relay channel* (TWRC) ao combinar princípios de *physical-layer network coding* (PLNC) com multiplexagem espacial com *multiple-input multi-output* (MIMO) e ainda com rádios usando *full duplex*. Esta combinação conduz a uma muito considerável redução do número de *time-slots* para trocar pacotes de mensagens entre todos os terminais de uma rede sem fios por meio de um retransmissor central.

Este trabalho começa por introduzir o conceito de PLNC e de reticulados (*lattices*). As estratégias de detecção mais importantes para a tradicional multiplexagem espacial MIMO são descritas e foram todas implementadas para comprovar a sua viabilidade para posterior concatenação com outras técnicas de processamento de sinal para supressão de interferência e auto-interferência.

Todas as estratégias propostas baseiam-se em *virtual MIMO* na fase de *uplink* usando detectores *lattice reduction aided* (LRA). Esta fase é sempre seguida por uma fase de *broadcast* para os terminais no *downlink*.

A dissertação aborda uma rede com três terminais com um retransmissor com múltiplas antenas no qual um esquema baseado em *time division multiple access* (TDMA) necessitaria de seis *time-slots* para trocar mensagens entre todos os terminais. Este trabalho mostra que, combinando certas estratégias de trocas de mensagens com processamento de sinal na camada física, é possível trocar todas as mensagens usando apenas três, dois, e, no regime em full-duplex, em apenas um *time-slot*, mantendo um desempenho semelhante desde que exista um bom controlo da auto-interferência.

Palavras-chave: TWRC, MIMO, PLNC, full duplex, LRA, TDMA, throughput

Abstract

This dissertation proposes new concepts beyond the traditional two-way relay channel (TWRC) combining principles of physical-layer network coding (PLNC) systems with multiple-input multi-output (MIMO) spatial multiplexing and full duplex radios, which lead to a considerable reduction in the number of time-slots necessary to exchange information between all terminals through a relay.

The research starts by making an introduction to the concept of PLNC and lattice and the currently most important detection strategies for "pure" MIMO are explained and implemented to assess their viability for later concatenation with other signal processing techniques to loop-interference supression.

All the proposed strategies use virtual MIMO on the uplink phase while using lattice reduction aided (LRA) detection. This phase is always followed by a broadcast phase in the downlink.

The dissertation assesses a three terminals network with a MIMO relay for which a time division multiple access (TDMA) based scheme would require 6 time-slots to exchange messages between all terminals. The conclusion is that a number of strategies and techniques eventually allow us to exchange all the messages using only three, two, and ultimately, using full-duplex, in just one time-slot, as long as the loop-interference is efficiently reduced.

Keywords: TWRC, MIMO, full duplex, PLNC, LRA, TDMA, throughput

Acknowledgments

First of all, I want to thank my Professor Francisco Monteiro for his supervision, and finding time for the countless hours spent on supervision and the many discussions we had that allowed me to improve my knowledge of this field.

A kind word to my colleague Flávio Brás for the many afternoons we spent brainstorming to find innovative ways to approach the problems.

Many thanks to my family, especially my parents and sister. Also my girlfriend for her huge support all these years we have spent together.

To Instituto de Telecomunicações (IT), for providing financial support during part of my master programme under the FCT project PEst-OE/EEI/LA0008/2013.

Finally, I also thank my friends, especially bNa for always being there for over a decade.

Index

List of Figures

List of Tables

Acronyms

BER	Bit error rate
BLAST	Bell Laboratories layered lpace-lime lrchitecture
CDMA	Code-division multiple access
CF	Compute-and-forward
CLLL	Complex LLL
CSIR	Channel state information at the receiver
CVP	Closest vector problem
DS	Direct sequence
EM	Electromagnetic
FD	Full duplex
FDMA	Frequency division multiple access
FH	Frequency hopping
GSM	Global system for mobile communications
KZ	Korkine-Zolotareff algorithm
LI	Loop-interference
LIC	Loop-inteference cancellation
LLL	Lenstra Lenstra Lovász algorithm
LST	Layered space-time
LTE	Long term evolution
MIMO	Multiple-input multiple-output
ML	Maximum likelihood

MMSE	Minimum mean square error
OFDM	Orthogonal frequency division multiplexing
OSIC	Ordered successive interference cancelation
PLNC	Physical-layer network coding
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase-shift keying
SDMA	Space-division multiple access
SIC	Successive interference cancelation
SNR	Signal-to-noise ratio
TDMA	Time division multiple access
TWRC	Two-way relay channel
V-BLAST	Vertical Bell Laboratories layered space-time architecture
WiMax	Worldwide interoperability for microwave access
ZF	Zero-forcing
ZMSW	Zero-mean spatial white

Symbols

$\mathbf{B}$	Lattice generating matrix
$\mathbf{B}_{\mathrm{red}}$	Reduced lattice matrix
C	Constellation
D	Diversity order
$\mathrm{E_s}$	Average symbol energy
$\mathrm{g_T}$	Transmitter filter
$\mathrm{g_R}$	Receiver filter
$\mathbf{H}$	Channel matrix
$\mathbf{H}^{\dagger}$	Pseudo-inverse matrix **H**
$\mathbf{H}^{\mathrm{H}}$	Hermitian matrix **H**
$\bar{\mathbf{H}}$	Extended matrix **H**
$\mathbf{H}_{\mathrm{red}}$	Reduced channel matrix
$\mathrm{h_{ij}}$	Channel fading coefficient
h^{-1}	Matched filter
$\mathrm{I_N}$	Identity matrix $N \times N$
k	Rank of the lattice
L	Number of terminals

M	Unimodular matrix
n	Additive noise vector
N_R	Number of receive antennas
N_T	Number of transmit antennas
T_s	Symbol transmission period
U	Unitary matrix
x	Transmitted symbol vector
$\hat{\mathbf{x}}$	Estimated symbol vector
W	Receiver filter
$\mathbf{w}_i$	Message from transmitter i
y	Received vector
$\bar{\mathbf{y}}$	Extended received vector
ϵ	Small real number
Λ	Lattice
σ_n^2	Noise variance
σ_x^2	Variance of constellation symbols
C^n	*n*-dimensional complex space
R^n	*n*-dimensional real space
Z^n	*n*-dimensional vectors with integer coordinates
$L(\mathbf{B})$	Lattice generated by basis **B**
$\det(\mathbf{B})$	Determinant of **B**

$(.)^H$ Hermitian operator

$(.)^T$ Transposition

$P(B)$ Fundamental region of a lattice

$Q_C[.]$ Quantization to constellation *C*

$\mathrm{vol}(\Lambda)$ Volume of a lattice

$V(\Lambda)$ Voronoi region of a lattice

$I(.)$ Imaginary part of a complex number

$R(.)$ Real part of a complex number

Chapter 1

Introduction

This chapter will introduce the basic concepts explored thru this dissertation, in special to Physical Layer Network Coding

Section 1.1 – Overview

In the past few years we witnessed a great increase in devices that utilize wireless networks to access multimedia applications. Since these devices consume a larger bandwidth (due to the higher data rates required), in the near future networks will be facing a bottleneck since the numbers of transmission frequencies is finite and interference between devices transmissions is a problem.

All wireless communications are transmitted in the form of electromagnetic (EM) waves, therefore all messages sent by a node will be received by all antennas nearby. Moreover, the receiver can be receiving multiple EM signal at the same time, and this superimposes interference to the received signal. Traditional communications systems tends to avoid or reduce the amount of interfering signals since packets collisions result in the loss of both packets.

Nowadays the existing communications systems share the channel through time divisions multiple access (TDMA), which uses the concept of time-slot (TS) in order to avoid packets collision. Other much utilized strategy is frequency division multiple access (FDMA) where each user is given a part of the spectrum to transmit the packets, normally FDMA is used to separate the uplink (communications from the user equipment (UE) to the base station (BS)) and the downlink (communications from the BS to the UE). The GSM standard is an example that combines these two strategies.

Today's wireless systems typically use orthogonal frequency division multiplexing (OFDM) as a modulation scheme, and long term evolution (LTE) is an example of how to combine TDMA and FDMA with OFDM.

As mentioned before the spectrum is limited so engineers needed to find new ways on how to increase spectral efficiency. A new concept appeared, space-division multiple access (SDMA), where these system use multiple space dimensions in order to transmit EM

signals; multiple-input multiple-output (MIMO) technologies takes full advantage of multiple space divisions.

MIMO was the first technology to uses channel interference as an advantage to the system since all the others consider channel interference as one important cause for information loss. A new idea was recently presented that also uses channel interference in order to increase channel capacity in wireless networks, physical-layer network coding (PLNC). Both these technologies (MIMO and PLNC) will be the main focus on this dissertation, as well as the decoding technics used.

Section 1.2 – Physical-layer network coding

In 2003 a new idea was developed by Li, Yeung and Cai [1], named Network Coding (NC). Three years later a new concept of NC was proposed that uses channel interference to the network advantage, this idea have been proposed by three independent research groups: Zhang, Liew and Lam [2] Popovki and Yomo [3]; Nazer and Gastpar [4]. This idea takes advantage of a naturally occurring NC in Nature, were EM waves sharing the same physical space add up, and this turn out to implement the same concept proposed in [1], however without the need for any computing of the bits at the relay.

To further exemplify the concept of NC and PLNC let us assume a very simple channel composed by two users that can't hear each other and a relay in the middle, a two-way relay channel (TWRC).Figure 1, it is assumed that all terminals share the same frequency band and they operate in half-duplex, were each terminal can only send or receive a message during a single TS. These concepts will be compared with a more traditional scheme such as TDMA.

Figure 1 – Two-Way relay channel.

Both users (user 1, user 2) want to transmit their messages (w_1, w_2) to each other. Using TDMA on the first TS user 1 send w1 to the relay, then on the second TS the relay broadcast w1 to user 2 and the process is repeated, third TS user 2 send w2 to the relay and on the fourth TS the relay broadcast w2 to user 1, as exemplified on Figure 2. So it takes four TS for the users to exchange messages.

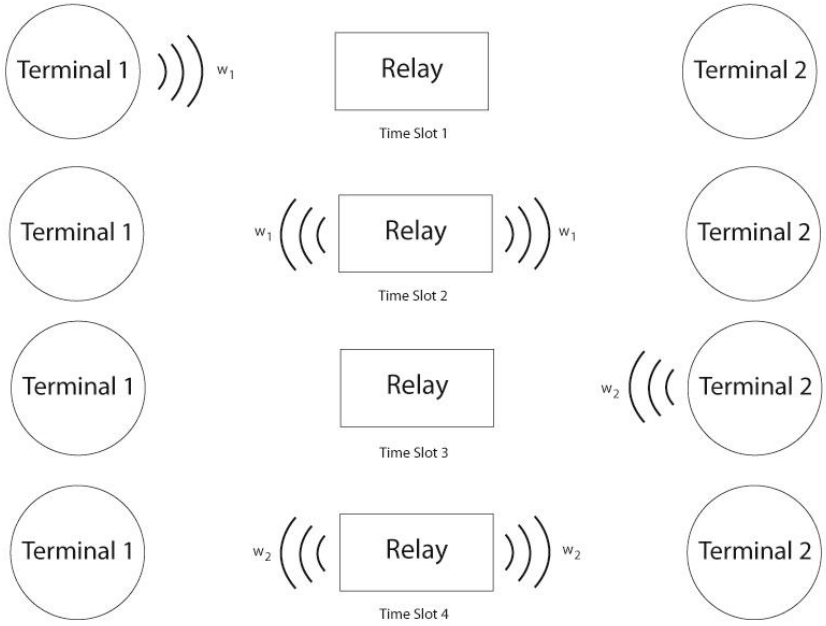

Figure 2 – TDMA.

Using NC on the first TS user 1 sends w1 to the relay then on the second TS user2 sends w_2 to the relay were both messages are summed bit by bit and the relay broadcast the sum on the third TS, as exemplified on Figure 3. Since user 1 knows w_1 it can retrieve w_2 from the sum received, the same for user 2 with w_1. This ways only takes 3 TS for the messages to be exchanged, so there is a 33% increased throughput over TDMA.

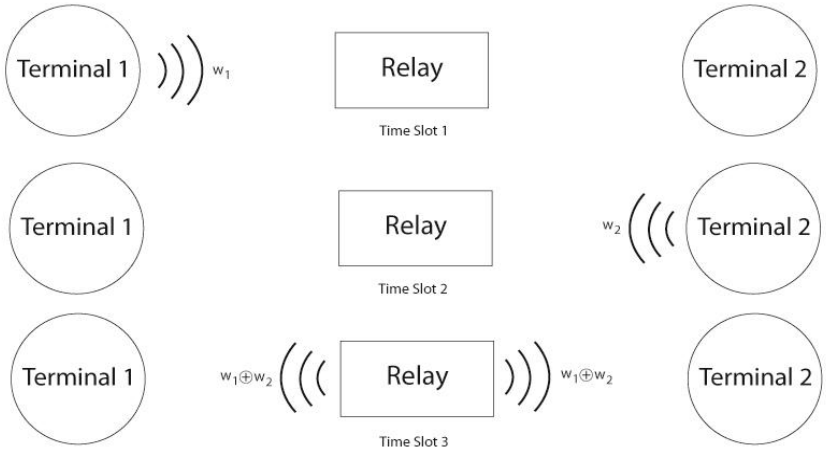

Figure 3 – Network Coding.

Using PLNC technics, on the first TS both users 1 and 2 send their messages, and as explained before, both EM waves will superimpose in the physical medium so the relay receives the sum of the two messages. On the second TS the relay broadcast the received sum to user 1 and user 2 Figure 4. The users are able to retrieve the messages in the same way that NC.

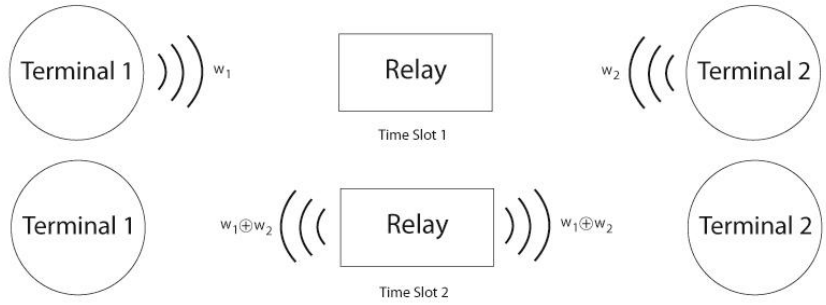

Figure 4 – Physical-Layer Network Coding.

In PLNC is very important to assume a carrier and symbol synchronization as well as power control so when the packets from both users arrive at the relay have the same phase and amplitude. Later on in this dissertation we will focus on the effect of fading channels and noise for PLNC model.

Most works on PLNC focus on the TWRC model but there is also some other important research being made on a multi-way relay channel where a relay or a system of

relays interconnects multiple end nodes. The work by Chen Feng, Danilo Silva and Frank Kschischang [5] developed seminal work on that.

Section 1.3 – Compute and forward

Nazer and Gatspar in [4] proposed a strategy called Compute-and-Forward (CF). It is a branch of PLNC were the relay decode linear equations of transmitted messages using linear combinations provided by the channel. If the destination gets sufficient linear combinations it can solve those linear equations obtaining the sent messages.

A compute-and-forward scheme is shown in Figure 5 were *L* terminals and a single relay are represented. Each of the terminals takes messages from a finite field, map them into lattice points and transmit them into the channel. The relay receives a linear combination of these points, as in Figure 6 and decode an integer combination of them, i.e., creates a Diophantine equation. This equation of lattice point is then mapped back to linear equations over a finite field. When the relay gets *L* different linear equation it can decode the desired messages [5].

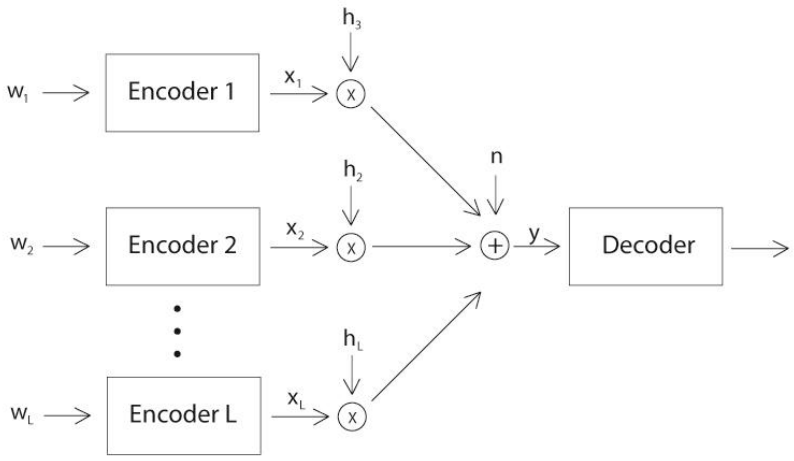

Figure 5 – Compute-and-Forward scheme.

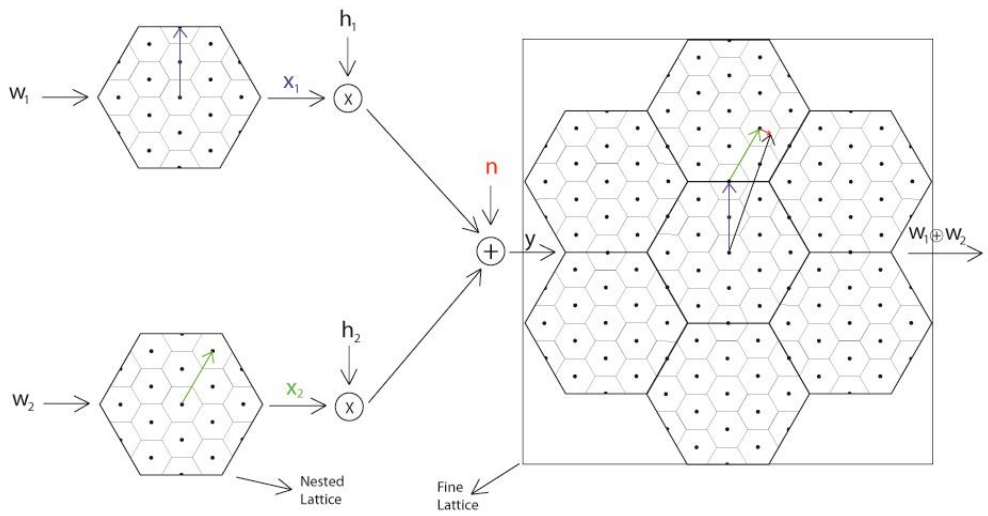

Figure 6 – Nested and Fine lattice in Compute-and-Forward.

With compute-and-forward the transmission gets is more protected against noise and raises the opportunity to exploiting interference for cooperative gains. This strategy relies on codes with a linear structure, specifically nested lattices codes. The lattice code should have some form of modulo arithmetic so that we can map between the linear combination taken by the channel and our desired combination over the messages. This property is satisfied by nested lattices which are a subset of another lattice, called the fine lattice. The nested lattice can be replicated tilling the entire fine lattice. Note that lattices will be presented in detail in Chapter 2, since nested lattices are not discussed Figure 6 illustrates a simple example of its concept through a system were two transmitters send a message at the same time and frequency to a receiver. Each transmitter maps its finite-field message into an element of the nested lattice code and sends this vector on the cannel. Here, the channel coefficients (h_1 and h_2) are taken to be equal to 1. Therefore, the receiver observes a noisy sum of the transmitted vectors and determines the closest lattice point. After taking a module operation with respect to the nested lattice, the receiver can invert the mapping and determine the module sum of the original messages [6].

Chapter 2

Lattices

Lattices are introduced in this Chapter providing basic and useful properties, to be used later on, in a more formal way

Section 2.1 – Introduction

Lattices are *discrete* regular arrangements of points in Euclidean space, or in other words, they are a set of points in n-dimensional space with a periodic structure, such as the ones illustrated in Figure 7. Three dimensional lattices occur naturally in many settings such as crystals or in simple stack of oranges. Historically, lattices were investigated since the late 18th century by mathematicians such as Lagrange, Gauss and later Minkowski, and was about the publication of the last one, Hermann Minkowski, called *Geometrie der Zahlen* in 1896 that lattices have become a standard tool in number theory, especially in the areas of algebraic number theory and the arithmetic theory of quadratic forms for instance. Going forward in time, a significant advance in the algorithmic theory of lattices of general rank occurred in 1982, with the development of the powerful lattice basis reduction algorithm that came to be called the LLL algorithm (named after the name of its inventors) that has found numerous applications in both pure and applied mathematics[8]. A complex-value version of this famous algorithm will be present later on in this chapter.

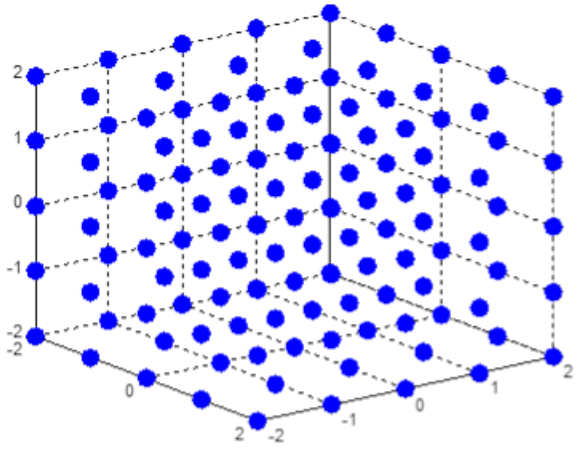

Figure 7 – Lattice illustrations in $\mathbb{R}^3$.

Lattices have now many applications in computer science and mathematics. Among many others applications they are helpful to the solution of integer programming problems, cryptanalysis, Diophantine approximation (finding rational approximations for real numbers) and the design of error correcting codes for multi-antenna systems. More

recently, one of the most promising applications and that has attracted much attention is the use of lattices as a source of computational hardness for the design of secure cryptographic functions.

Another promising application of lattices is its use as base of coding schemes and that application is the one that will be explored. In the present work lattices take a prominent place as it will become clearer in the next chapters so it is important to define this structure and understand its main properties.

Section 2.2 – Basis definitions

In this subsection it will be presented a more formal definition of lattice as well as its most important mathematical properties based on [9] and [10].

There are several ways of specifying a lattice. One of the definitions says that a lattice is a discrete additive subgroup of $\mathbb{R}^n$, i.e., it is a subset $\Lambda \subseteq \mathbb{R}^n$ satisfying the following properties:

- (subgroup) Λ is closed under addition and subtraction,
- (discrete) There is an $\epsilon > 0$ such that any two distinct lattice points $x \neq y \in \Lambda$ are at distance at least $\|x - y\| \geq \epsilon$.

Note that not every subgroup of $\mathbb{R}^n$ is a lattice. For example the subgroup $\mathbb{Q}^n$ is not a lattice because it not fulfils the second property while the subgroup $\mathbb{Z}^n$ is a lattice because integer vectors can be added and subtracted obtaining again an integer vector, and clearly the distance between any two integer vectors is at least one.

An equivalent definition of lattices can be obtained from $\mathbb{Z}^n$ by applying a linear transformation. To illustrate this definition let us consider a matrix $\mathbf{B} \in \mathbb{R}^{n \times k}$ that has full column rank, what means that the columns of this matrix are linearly independent, and then $\mathbf{B}(\mathbb{Z}^n) = \{\mathbf{Bx}: \mathbf{x} \in \mathbb{Z}^n\}$ is also a lattice. Obviously this set is closed under addition and

subtraction, besides it is also discrete. Moreover all lattices can be expressed as $\mathbf{B}(\mathbb{Z}^n)$ for some B what leads to the following definition.

Let $\mathbf{B} = [b_1, \dots, b_k] \in \mathbb{R}^{n \times k}$ be a linearly independent vectors in $\mathbb{R}^n$. The lattice generated by B is the set of all the integer linear combinations of the columns of B.

$$\mathrm{L}\left(\mathbf{B}\right) = \left\{\mathbf{B}\boldsymbol{x} : \mathrm{x} \in \mathrm{Z}^{\mathrm{k}}\right\} = \left\{\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{x}_{\mathrm{i}}.\mathrm{b}_{\mathrm{i}} : \mathrm{x}_{\mathrm{i}} \in \mathrm{Z}\right\} \quad (1)$$

The matrix $\mathbf{B}$ is called a basis for the lattice $\mathcal{L}(\mathbf{B})$. The integer n is called the dimension of the lattice; the integer k is the rank of the lattice and in the case of $n = k$, $\mathcal{L}(\mathbf{B})$ is called a full rank lattice.

In this second definition lattices can be represented by a basis matrix $\mathbf{B}$, containing integer or rational entries, which generate the lattice. It is important keep in mind this definition because it will be present during all the work. The definition can be extended to complex lattices however it is possible to transform any complex lattice into a real lattice through the real equivalent model, so the descriptions about lattices will continue with real lattices.

A lattice is the span of a finite set of vectors in a Euclidean space:

$$\Lambda = \mathcal{L}(\mathbf{B}) = \{\mathbf{Bx} | \mathbf{x} \in \mathbb{Z}^k\}. \quad (2)$$

Notice the similarity between the definitions of a lattice and the span of a set of vectors B:

$$span(B) = \{\mathbf{By}: \mathbf{y} \in \mathbb{R}^k\}. \quad (3)$$

The central difference is that in a lattice *only* integer coefficients are allowed, resulting in a discrete set of points. As vectors $b_1, \dots, b_n$ are linear independent, any point $y \in span(B)$ can be written as a linear combination $\mathbf{y} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$ in a unique way. Therefore $y \in \mathcal{L}(\mathbf{B})$ if and only if $x_1, \dots, x_n \in \mathbb{Z}$.

Notice that the definition $\mathcal{L}(\mathbf{B}) = \{\mathbf{Bx}: \mathbf{x} \in \mathbb{Z}^m\}$ can be extended to matrices B whose columns are not linearly independent. However, in this case, the resulting set of points is not always a lattice because it may not be discrete.

In the case of full-rank lattices the volume of the lattice (that it is the same of the volume of any of its many possible fundamental regions) is

$$vol(\Lambda) = |\det(\mathbf{B})|, \tag{4}$$

however, in general the following expression is required:

$$vol(\Lambda) = \sqrt{det(\mathbf{B}^T\mathbf{B})} \tag{5}$$

In the complex case the Hermitian operator replaces transposition in the equation above. The volume of the lattice is an invariant of the lattice, i.e., is independent of the choice of basis.

Let $\Lambda = \mathcal{L}(\mathbf{B})$ be a lattice of rank n. The determinant of Λ, denoted $det(\Lambda)$, as the n-dimensional volume of the fundamental region $P(\mathbf{B})$. In symbols, this can be written as $det(\Lambda) := \sqrt{det(\mathbf{B}^T\mathbf{B})}$. In the special case that Λ is a full rank lattice, $\mathbf{B}$ is a square matrix, and it stays that $det(\Lambda) = |\det(\mathbf{B})|$.

The determinant of a lattice is well-defined, in the sense that it is also independent of the choice of basis $\mathbf{B}$. Indeed, if $\mathbf{B}_1$ and $\mathbf{B}_2$ are two bases of Λ, then as it will be present next $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$ for some unimodular matrix $\underline{\mathbf{U}}$. Hence,

$$\sqrt{\det(\mathbf{B}_2^T\mathbf{B}_2)} = \sqrt{\det(\mathbf{U}^\mathbf{T}\mathbf{B}_1^T\mathbf{B}_1\mathbf{U})} = \sqrt{\det(\mathbf{B}_1^T\mathbf{B}_1)} \tag{6}$$

The determinant of a lattice is inverse proportional to its density, what means that the smaller the determinant, the denser the lattice is.

In the section 2.3 named Fundamental Region it was presented a lemma to verify if a basis can be a basis of a certain lattice based on the fundamental region that the vectors of the basis create. Now let present it in a more formal way.

Let Λ be a lattice of rank n, and let $b_1, \dots, b_n \in \Lambda$ be n linearly independent lattice vectors. Then $b_1, \dots, b_n$ form a basis of Λ if and only if $P(b_1, \dots, b_n) \cap \Lambda = \{0\}$.

Proof. Assume first that $b_1, \dots, b_n$ form a basis of Λ. Then, by definition, Λ is the set of all their integer combinations. Since $P(b_1, \dots, b_n)$ is defined as the set of linear combinations of $b_1, \dots, b_n$ with coefficients in the range $[0,1[$, the intersection of the two sets is $\{0\}$

For the other direction, assume that $P(b_1, \dots, b_n) \cap \Lambda = \{0\}$. Since Λ is a rank n lattice and $b_1, \dots, b_n$ are linearly independent, we can write any lattice vector $x \in \Lambda$ as $\sum y_i b_i$ for some $y_i \in \mathbb{R}$. Since by definition a lattice is closed under addition, the vector $x' = \sum(y_i - \lfloor y_i \rfloor)b_i$ is also in Λ. By our assumption, $x' = 0$. This implies that all y_i are integers and hence x is an integer combination of $b_1, \dots, b_n$.

Note that different set of vectors may generate the same lattice but all these different bases are related by unimodular transformations as it will be described below.

Let B_1 and B_2 be two bases. Then $\mathcal{L}(\mathbf{B}_1) = \mathcal{L}(\mathbf{B_2})$ if and only if there exists a unimodular matrix U such that $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$.

A matrix $\mathbf{U} \in \mathbb{Z}^{n \times n}$ is called unimodular if it is a square matrix with integer entries and determinant ± 1. The inverse of a unimodular matrix is also unimodular.

Proof. First assume $\mathbf{B}_1 = \mathbf{B}_2\mathbf{U}$ for some unimodular matrix U. Notice that if U is unimodular, then U^{-1} is also unimodular. In particular, both U and U^{-1} are integer matrices, and $B_1 = B_2U$ and $B_2 = B_1U^{-1}$. It follows that $\mathcal{L}(B_1) \subseteq \mathcal{L}(B_2)$ and $\mathcal{L}(B_2) \subseteq \mathcal{L}(B_1)$, i.e., the two bases B_1 and B_2 are equivalent and they generate the same lattice.

Now assume B_1 and B_2 are two bases for the same lattice $\mathcal{L}(B_1) = \mathcal{L}(B_2)$. Then, by definition of lattice, there exist integer square matrices V and W such that $B_1 = B_2W$ and $B_2 = B_1V$. Combining these two equations appears $B_1 = B_1VW$, or equivalently, $B_1(I - VW) = 0$. Since vectors B_1 are linearly independent, it must be $I - VW = 0$, i.e., $VW = I$. In particular, $\det(V) \cdot \det(W) = \det(V \cdot W) = \det(I) = 1$. Since matrices V and W have integer entries, $\det(V), det(W) \in \mathbb{Z}$, and it must be $\det(V) = \det(W) = \pm 1$.

A simple way to obtain a basis of a lattice from another is to apply (a sequence of) elementary column operations, as defined bellow. It is easy to see that elementary column operations do not change the lattice generated by the basis because they can expressed as right multiplication by a unimodular matrix. Elementary column operations are:

- Swap the order of two columns in B_1.
- Multiply a column by -1.
- Add an integer multiple of a column to another column: $b_i \leftarrow b_i\ a \cdot b_j$ where $i \neq j$ and $a \in \mathbb{Z}$.

Moreover, any unimodular transformation can be expressed as a sequence of elementary integer column operations. All these properties are essential in any lattice reduction algorithm when finding an equivalent basis to replace some other given one.

To further illustrate the definitions it will be shown some examples. In the following examples the fulfilled points are the ones that belong to the lattice. Besides these examples only contains a part of the lattice but is intuitive imagine how they span to the remaining space. Figure 8 a) represents a lattice generated by the vectors $(1,0)^T$ and $(0,1)^T$ which is the lattice of all integers' points, $\mathbb{Z}^2$. As it is possible to see in Figure 8 b) the basis $(1,1)^T$ and $(2,1)^T$ also generate $\mathbb{Z}^2$ moreover these bases are not unique, fact that will be explained later on in this chapter on section Bases. On the other hand Figure 8 c) des not generate $\mathbb{Z}^2$, as basis $(1,1)^T$ and $(2,0)^T$ generates instead a lattice of all integer points whose coordinates sum to an even number (called the *checkerboard lattice*). Apart from the other examples that are full-rank lattices Figure 8 d) represents a lattice of dimension 2 and rank 1 generated by the base $(2,1)^T$.

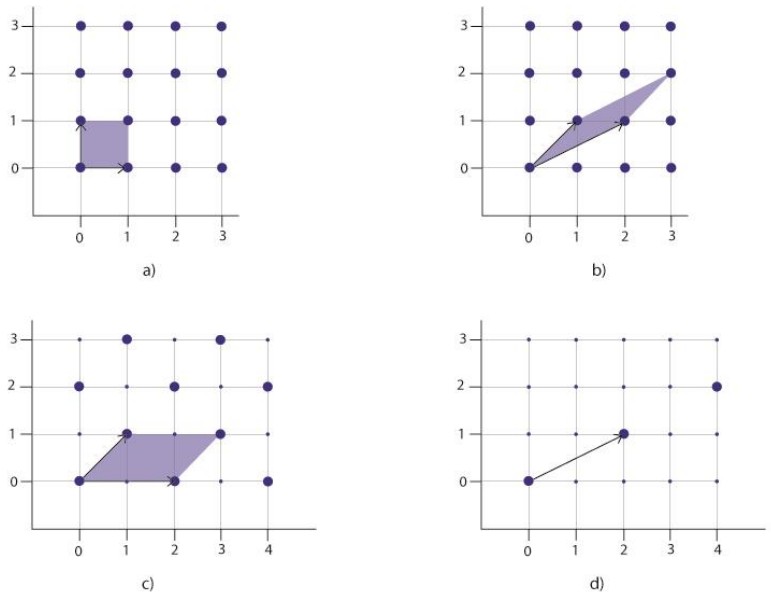

Figure 8 – Examples of lattices in $\mathbb{R}^2$. a) basis (1,0) and (0,1). b) basis (1,1) and (2,1). c) basis (1,1) and (2,0). d) basis (2,1).

Section 2.3 – Fundamental Region

For any lattice basis **B** the fundamental region or parallelepiped is defined as

$$P(\mathbf{B}) = \{\mathbf{Bx} | \mathbf{x} \in \mathbb{R}^n, \forall i: 0 \leq x_i < 1\}. \quad (7)$$

Notice that $P(\mathbf{B})$ depends on the basis **B** and it is easy to imagine that if we place one copy of $P(\mathbf{B})$ at each lattice point in $\mathcal{L}(B)$ we obtain a tiling of the entire $span(\mathcal{L}(\mathbf{B}))$ as shown in Figure 9.

In Figure 8 and Figure 9 the fundamental regions are represented by the gray areas. In Figure 9 are used two different gray tones just for a better illustration of the tilling of the space but they represent the same area, i.e. the same fundamental region.

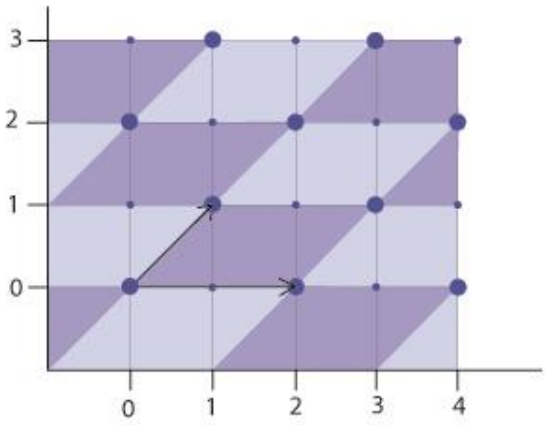

Figure 9 – Tilling of the span(L(B)) with the fundamental region P(B).

Base on the definition of the fundamental region it is possible to determine if a set of vectors forms a basis of a lattice. As it shown above in the examples, not every set of n linearly vectors in $\mathbb{Z}^n$ is a basis of $\mathbb{Z}^n$ and this is possible to determine following the lemma that says that the fundamental region generated by the vectors should not contain any lattice points, except the origin. For example, notice that the fundamental regions in Figure 8 a) and Figure 8 b) do not contain any nonzero lattice points and so they are bases of $\mathbb{Z}^2$, however the fundamental region of Figure 8 c) contains the lattice point $(1,0)$ and so the vectors that generate this lattice are not a basis of $\mathbb{Z}^2$.

Section 2.4 – Voronoi Region

The Voronoi region is defined by

$$v(\Lambda) = \{z \in span(\mathbf{B}): \|\mathbf{x} - \mathbf{z}\| < \|\mathbf{y} - \mathbf{z}\|, \forall \boldsymbol{y} \in \Lambda\}. \quad (8)$$

This equations defines the Voronoi region, which consist on the space where the lattice exists that contains all the points in the span of the lattice which are closer to a given lattice point x than to any other point in the lattice. This region is a characteristic of the lattice and independent of any particular generating matrix and it is the most interesting

fundamental region that tiles the entire space once it constitutes the optimal decision region for the closest vector problem (CVP) in a lattice, a problem that will be focused on the next chapter.

Figure 10 shows the tilling of two distinct lattices with its respective Voronoi region. The Voronoi regions of different lattices are always different (interestingly, it is known that in two dimensional case, the maximum number of facets they can have is six).

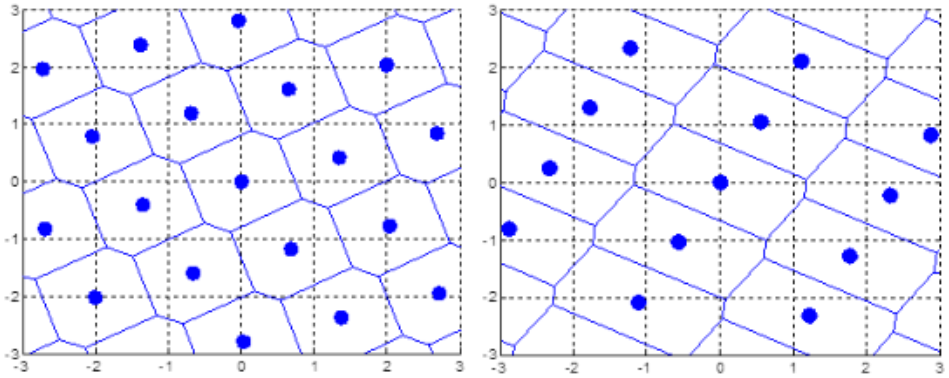

Figure 10 – Illustrations of Voronoi regions of two distinct lattices.

Section 2.5 – Successive Minima and Shortest Vector

One basic parameter of a lattice is the length, meaning the Euclidean norm, of the shortest nonzero vector in the lattice. This parameter is denoted by λ_1.

An equivalent way to define λ_1 is the following. It is the smallest r such that the lattice points inside a ball of radius r span a space of dimension 1 as is shown in Figure 11. This definition leads to the following generalization of λ_1 known as successive minima.

Let Λ be a lattice of rank n. For $i \in \{1, \dots, n\}$ it is define the ith successive minimum as

$$\lambda_i(\Lambda) = \inf\left\{r | \dim\left(\operatorname{span}\left(\Lambda \cap \overline{\mathbf{B}}(0, r)\right)\right) \geq i\right\} \tag{9}$$

where $\overline{\mathbf{B}}(0, r) = \{\mathbf{x} \in \mathrm{R}^m | \|\mathbf{x}\| \leq r\}$ is the closed ball of radius r around 0.

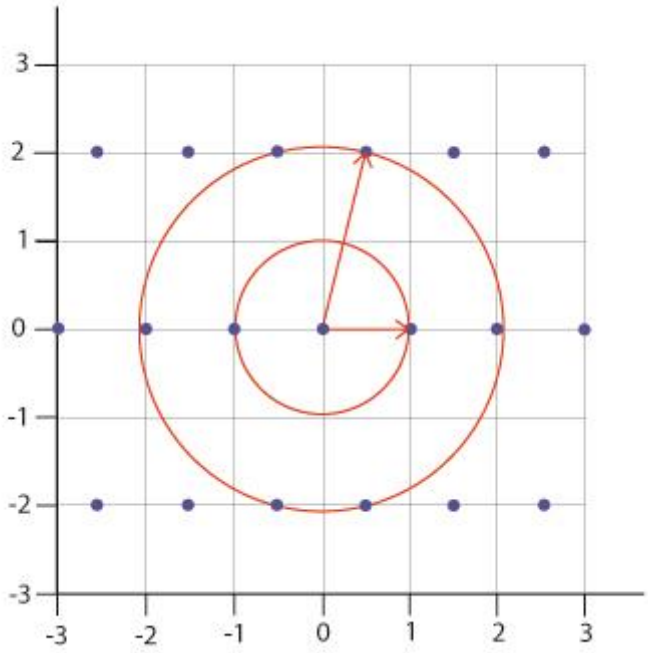

Figure 11 – Shortest vector.

Section 2.6 – Lattice reduction

The idea of lattice reduction consists in changing a basis B of a lattice Λ into a shorter basis $\mathbf{B}_{red}$ such that Λ remains the same. This process can be used to solve the shortest vector problem however for high rank basis there is no known algorithm that finds the shortest vector in polynomial time.

Lattice reduction can be implemented by algorithms such as LLL and KZ being the former the most known. The LLL lattice basis reduction algorithm was invented by Arjen Lenstra, Hendrik Lenstra and László Lovász in 1982 [11] and is a polynomial time lattice reduction algorithm. LLL usually obtains an approximation for the shortest vector but as it was mentioned there is not any efficient algorithm to solve this problem anyway the approximation obtain by the LLL suffices for many applications. In this dissertation a complex version of the original LLL algorithm (CLLL) will be used to assist a detection process in a receiver.

In sub-section 2.2 were presented simple techniques to obtain different bases that generate the same lattice, where one of those is swap the order of two vectors in **B** which is equivalent to apply a unimodular transformation. Roughly speaking what LLL algorithm does is perform successive orthogonal projections and if necessary uses the technique of swapping two consecutives vectors of **B**, in order to get a reduced or near orthogonal basis $\mathbf{B}_{\text{red}}$. So the output of the algorithm is a new basis $\mathbf{B}_{\text{red}}$ consisting of near-orthogonal vectors and a unimodular matrix **M** such that

$$\mathbf{B}_{\text{red}} = \mathbf{BM}. \tag{10}$$

In this dissertation, as mentioned, based on [12] a complex version of this algorithm was adopted once it reduces the complexity of the implementation without sacrificing any performance. The detailed pseudo-code, using Matlab notation, of the algorithm implemented is presented in Table 1. Notice that parameter δ controls the performance and the complexity of the algorithm, i.e., higher values of δ corresponds to higher complexity leading to a better performance. At the simulations performed along this dissertation δ was assumed to be 0.75 as advised in literature. Since in the following chapters basis **B** of the lattice will be considered as the matrix of the channel coefficients H the pseudo-code will use the latter nomenclature.

Table 1- Pseudo-code for the complex LLL algorithm

Input: H

Output: M and H_{red}

$[Q, R] =$ QR decomposition (H)
$\delta = 0{,}75$
$m = \text{size}(H, 2)$
$M = I_m$
$k = 2$
while $k \leq m$
    for $n = k - 1 : -1 : 1$
        $u = \text{round}((R(n,k)/R(n,n))$
        if $u \sim = 0$
            $R(1:n,k) = R(1:n,k) - u \cdot R(1:n,n)$
            $M(:,k) = M(:,k) - u \cdot M(:,n)$
        end
    end
    if $\delta|R(k-1,k-1)|^2 > |R(k,k)|^2 + |R(k-1,k)|^2$
        swap the $(k-1)^{th}$ and k^{th} columns in R and M
        $\Theta = \begin{bmatrix} \alpha^* & \beta \\ -\beta & \alpha \end{bmatrix}$ where $\alpha = \frac{R(k-1,k-1)}{\|R(k-1:k,k-1)\|}$ and $\beta = \frac{R(k,k-1)}{\|R(k-1:k,k-1)\|}$
        $R(k-1:k, k-1:m) = \Theta R(k-1:k, k-1:m)$
        $Q(:, k-1, k) = Q(:, k-1:k)\Theta^H$
        $k = \max(k-1, 2)$
    else
        $k = k + 1$
    end
end
$H_{red} = QR$

Chapter 3

MIMO detection

This Chapter will present MIMO detection problem as a closest vector problem in a lattice. Several detectors will be described and analyzed assessing theirs performances

Section 3.1 – MIMO spatial multiplexing

In recent years Multiple-input Multiple-output (MIMO) systems have emerged as one of the most promising technology to face the challenges described in Chapter 1.

MIMO communications systems can be described as using multiple antennas both on the transmitting and the receiving node. The main idea behind in MIMO is to use multiple parallel spatial "data pipes", increasing the data rate and adding diversity in order to improve quality of channel, by reducing the bit-error rate.

The importance of MIMO can be observed in many recent telecommunications standards, in combination with orthogonal frequency division multiplexing (OFDM) it is the base for the physical layer of the fourth generation (4G) of wireless networks. LTE and LTE-advanced, the most recent generation in Wi-Fi (IEEE 802.11.n), Worldwide interoperability for microwave access or WiMAX (IEEE 802.16m) are all standards that uses MIMO with OFDM.

The benefits for multiple antennas come from the use of spatial dimensions as a complement to time, since MIMO technology is also known as space-time wireless or smart antennas. The use of antennas arrays at one node is not a new concept, until the 1990s this idea was mainly used for beamforming, with increase the signal-to-noise ratio (SNR) by focusing into the desired direction, and spatial diversity, were in the presence of multipath propagation the SNR can be significantly improved by combining the output of uncorrelated antenna elements.

New proposals for using antennas arrays were introduced in the mid 1990s to increase the capacity of wireless channels beyond diversity, since diversity is only the first step to multipath propagation. When MIMO systems were introduced multipath was effectively converted for the benefit of wireless communications systems, because it takes advantage of random fading to multiply transfer rates [13].

Paulraj and Kailath have patented on of the first practical applications for MIMO [14] were they introduce a technique for increasing the channel capacity by using multiple antennas at both ends.

Foschini introduces the concept of Layered space-time (LST) architecture [15] in 1996, later this was known as Bell Laboratories Layered Space-Time Architecture (BLAST) and it was designed for a point-to-point MIMO communications system. To solve some of the limitations of BLAST a new version was proposed in [16] called Vertical Bell Laboratories Layered Space-Time Architecture (V-BLAST). With this types of architecture was possible, under certain fading conditions, transmit independent data stream over the matrix channel **H** [17]. A large number of papers started to appear based on Foschini proposal, then Viswanath, Tse and Anantharam were among the first contributions to MIMO multiple-access channels [18].

Until the year 2000 MIMO investigation only focus on single user applications, then MIMO started to expand to multiuser (multiple receivers) and network (multiple transmitters) applications, with new ideas new challenges appears as coding channel models and decoding techniques [19]. In recent years the main focus of the research community is on large and massive MIMO, where the number of antennas is bigger than 8. This new approach uses compact antennas and claims it can support a better performance while allowing fast iterative receiver decoding [20].

MIMO spatial multiplexing has allowed unprecedented spectral efficiencies in wireless fading channels achieving high data-rates. This gain in performance comes with high complexity detection in the receivers which poses a challenge in implementation. In this chapter is going to be presented most of the receivers that have been implemented. The performance of linear receivers, successive interference cancellation (SIC) receiver and lattice reduction aided receivers will be analyzed.

Thru this chapter will be used a point-to-point communication system using MIMO technology, as shown in Figure 12. This system has multiple antennas on both the transmitter and the receiver with a possible implementation in semi-mobile local-area

wireless data communications, for example a laptop computer with multiple antennas communicating with an access point, also equipped with multiple antennas.

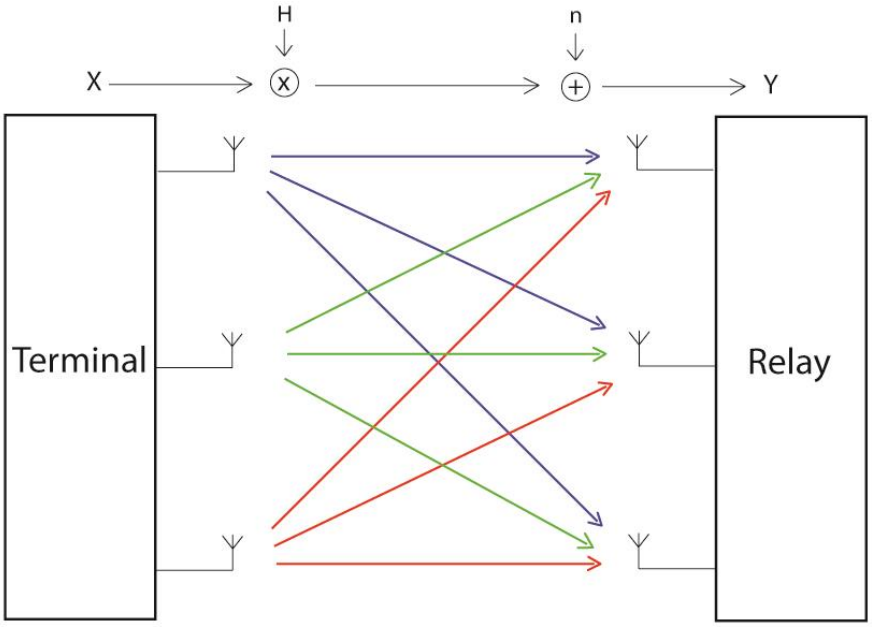

Figure 12 – MIMO system.

In this MIMO communications system with N_T transmit antennas and N_R receive antennas (with $N_R \geq N_T$ so that the linear system it gives rise to is determined) the relation between the transmitted and received signals can be modelled in the baseband as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{11}$$

where $\mathbf{y} = [y_1, \ldots, y_{N_R}]^T \in C^{N_R \times 1}$ is the received signals vector and $\mathbf{x} = [x_1, \ldots, x_{N_T}]^T \in C^{N_T \times 1}$ is the transmitted signals vector. The radio links between each pair of transmit and receive antennas are represented by the channel matrix $\mathbf{H} \in C^{N_R \times N_T}$ in which its entries h_{ij} represent the complex coefficient associated with the link between the pair of a i^{th} receive antenna and the j^{th} transmit antenna. Each h_{ij} is taken from a zero-mean circularly symmetric complex Gaussian distribution with unit variance, which corresponds to having a variance equal to $1/2$ in both real and imaginary components. In order to have an independent and identically distributed Rayleigh fading channel model the phase of each entry h_{ij} is uniformly distributed in $[0\ 2\pi[$ and their amplitude has a Rayleigh distribution.

In this model the vector $\mathbf{n} = \left[n_1, \ldots, n_{N_R}\right]^{\mathrm{T}} \in \mathrm{C}^{N_R \times 1}$ represents the noise vector that is added to the incoming signal vector. The entries of n are random variables taken from an independent circularly symmetric complex Gaussian with zero average and variance σ_n^2 , so that both its real and imaginary components have variance $\sigma_n^2/2$. This noise model is usually called as zero-mean spatially white (ZMSW) noise [21].

Full channel knowledge at the receiver is required in the model assumed in this chapter, usually known as channel state information at the receiver (CSIR). However, acquiring this knowledge is not a straightforward task, especially in rapidly time-varying channels, but this matter is beyond the scope of this work.

Throughout this dissertation, square quadrature amplitude modulation (QAM) constellations are assumed. Specifically it will be considered M-QAM constellations with $M = 4, 16, 64$ Figure 13 and the input symbols in each transmit antenna are taken from a finite complex constellation C constructed from the Cartesian product $C = C_{\mathrm{R}} \times C_{\mathrm{R}}$, where $\mathrm{C_R}$ is the real alphabet

$$\mathrm{C} = \left\{-\left(\sqrt{M} - 1\right), \ldots, -3, -1, +1, +3, \ldots, +\left(\sqrt{M} - 1\right)\right\}. \tag{12}$$

The average energy of the complex symbol taken from C is given by

$$E_s = \frac{1}{M}\sum_{\mathrm{x_i} \in \mathrm{C}} |x_{\mathrm{i}}|^2 \tag{13}$$

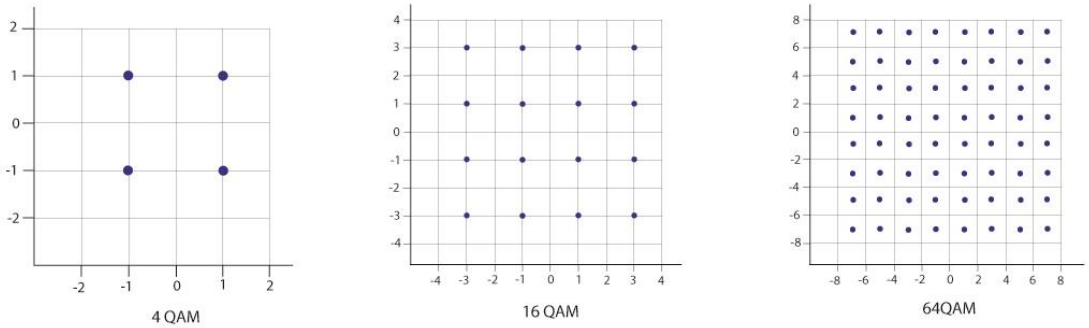

Figure 13 – M-QAM constellations.

assuming, without loss of generality, that the filters at the receiver have impulse response $h(t)$ normalized to $\int |h(\mathrm{t})|^2 \ \mathrm{dt} = 1$.

Considering that each y_i receives the sum of N_T symbols weighted by unit power random variables, i.e., $E\left[\left|h_{ij}\right|^2\right] = 1$, on *average* it is valid to calculate the SNR at the receiver as

$$\frac{E\{\|y\|^2\}}{E\{\|n\|^2\}} = \frac{E\{\|Hx\|^2\}}{E\{\|n\|^2\}} = \frac{E\left[\sum_{i=1}^{N_R}\sum_{j=1}^{N_T}\left|h_{ij}x_j\right|^2\right]}{E\left[\sum_{i=1}^{N_R} n_i^2\right]} = \frac{N_T N_R \sigma_x^2}{N_R \sigma_n^2} = N_T \frac{\sigma_x^2}{\sigma_n^2}. \quad (14)$$

Throughout this dissertation the performance results will be plotted as symbol error rate (SER) as function of the SNR.

In order to evaluate the performance of the receivers it will be considered the diversity order, d, or simply the slope. This parameter describes how the SER decreases with the increase of SNR, i.e., when plotting the SER as a function of the SNR this diversity order is simply the slope of the SER curve and it can be obtain by

$$d = \frac{\log(SER)}{\log(SNR)}. \quad (15)$$

Section 3.2 – The closest vector problem

One of the main problems considered in this thesis is the detection of a vector **x** given the noisy observation **y**. This topic has been a main researching problem in spatial multiplexing and assuming that all vectors **x** are equiprobable this detection can be seen as maximum likelihood (ML) detection which will be analyzed later on. In lattice theory this problem is known as closet vector problem (CVP).

The CVP can be described as the problem of finding the **x** that better explains the observation **y**, which is the one that after the linear transformation (**Hx**) generates the closest vector to the received vector **y**, i.e., it corresponds to the application of the maximum likelihood principle. Figure 14 exemplifies this problem in a simply case in $\mathbb{Z}^2$.

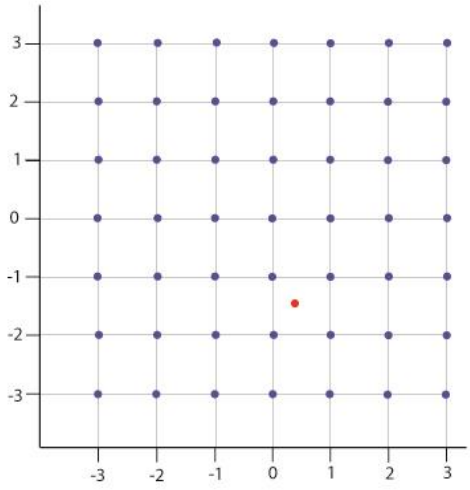

Figure 14 – Illustration of a example of CVP in Z^2.

CVP, from an algorithmic complexity point of view, is proven to be NP-hard, which is the worst case scenario in the hierarchy of complexity classes [22]. However, this problem can be solved approximately by means of a number of sub-optimal techniques such as zero-forcing (ZF) and successive interference cancelation (SIC), proposed by Babai in [23], or techniques based on lattice reduction [24]. This last technique offers some benefits: in quasi-static fading channels the complexity of lattice reduction is negligible because for a long frame of data the channel remains unchanged and it attains near-optimum performance. In addition, as it was discussed on the previous chapter, there are reduction algorithms such as LLL (or its CLLL, the complex counterpart) algorithm with polynomial complexity and with complexity independent of SNR [25]. These characteristics make lattice-reduction-aided (LRA) decoding especially suited for MIMO communications.

The following sub-sections will introduce the most important types of MIMO receivers. After a brief overview their performance will be shown with different number of antennas, keeping $N_T = N_R = N$, and with different M-ary QAM modulations. Specifically, it will be analyzed the 2×2 ($N = 2$) and 3×3 ($N = 3$) MIMO systems. The implemented model is the one previously described and the performance of each receiver will be assessed by plotting the SER as function of the SNR.

The simplest receivers will be introduced first, linear receivers such as zero-forcing (ZF) and minimum mean square error (MMSE), followed by the ordered successive interference cancelation (OSIC) algorithm, and finally the receivers using lattice reduction-aided (LRA) approach will be brought into discussion.

The performance results shown in this chapter serve to demonstrate that the system model and all the types of receivers considered are well calibrated and equal the results available in the literature.

Section 3.3 – Maximum Likelihood Detection

Assuming the zero-mean spatially white (ZMSW) noise model introduced in the beginning of this chapter, the probability density function of **y** given **H** and **x** can be written as

$$\mathrm{p}(\mathbf{y}|\mathbf{H},\mathbf{x}) = \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^{\mathrm{N}/2}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{2\sigma_{\mathrm{n}}^2}\right) \tag{16}$$

and consequently the maximum likelihood estimate ($\hat{\mathbf{x}}$) [26] for **x** given **y** is

$$\begin{aligned} \hat{\mathbf{x}}_{\mathrm{ML}} &= \arg\max_{\mathrm{x}\in\mathrm{C}^{\mathrm{N}}} \mathrm{p}(\mathbf{y}|\mathbf{H},\mathbf{x}) \\ &= \arg\max_{\mathrm{x}\in\mathrm{C}^{\mathrm{N}}} \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^{\mathrm{N}/2}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{2\sigma_{\mathrm{n}}^2}\right) \\ \hat{\mathbf{x}}_{\mathrm{ML}} &= \arg\min_{\mathrm{x}\in\mathrm{C}^{\mathrm{N}}} \|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2 \end{aligned} \tag{17}$$

The exponential growth of the search space for M-QAM constellations and the increase of dimensions N discourage the use of brute force maximum-likelihood detection in many practical systems given that simply evaluating $\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2$ for all possible $\mathrm{x} \in \mathrm{C}^{\mathrm{N}}$ requires too much time consumption.

The simulation setup implemented is designed to support any number of antennas and M-QAM constellations with $M = 4, 16, 64$. However due to computational constraints

in this dissertation will only use 2×2 and 3×3 MIMO configurations as well as just 4 and 16-QAM constellations will be used to show the performance of the receivers.

In Figure 15 and Figure 16 is shown the performance of maximum likelihood (ML) detection for 2×2 and 3×3 MIMO configurations. This strategy achieves diversity order equal to N as proven in literature, which means that it captures all the spatial diversity for spatial multiplexing (although not the one attained by space-time codes!), and so the ML performance will be always depicted in the following results along this chapter as a comparison benchmark for the best possible attainable performance. Along them its performance was simulated for each specific case so there are some variations of its performance caused by the variations of the specific channels considered. Notice that, as expected, 16-QAM constellation suffers from a power penalty of some dB, i.e., it achieves the same SER value from the 4-QAM constellation only in a higher value of SNR.

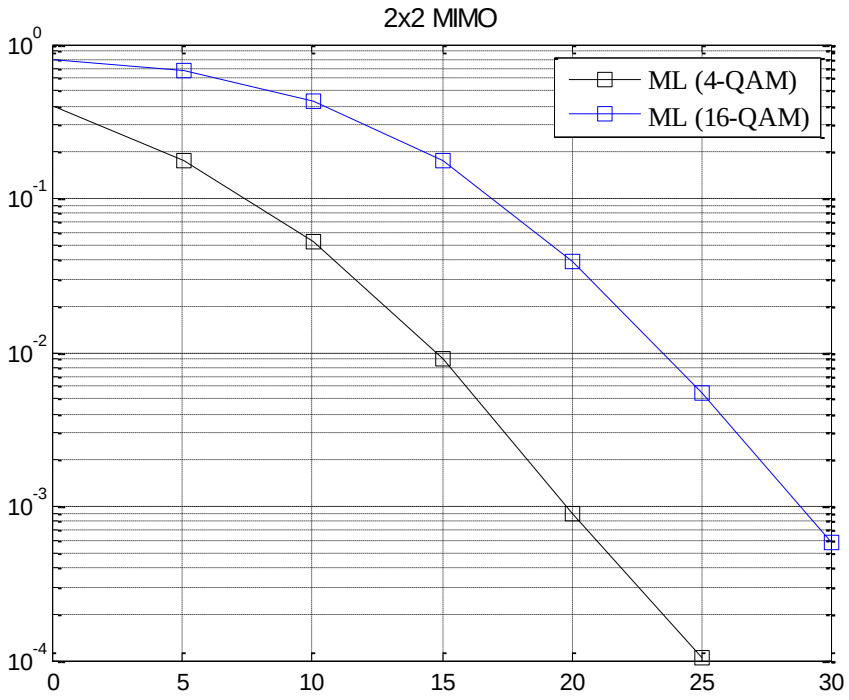

Figure 15 – ML detection for 2 x 2 antennas.

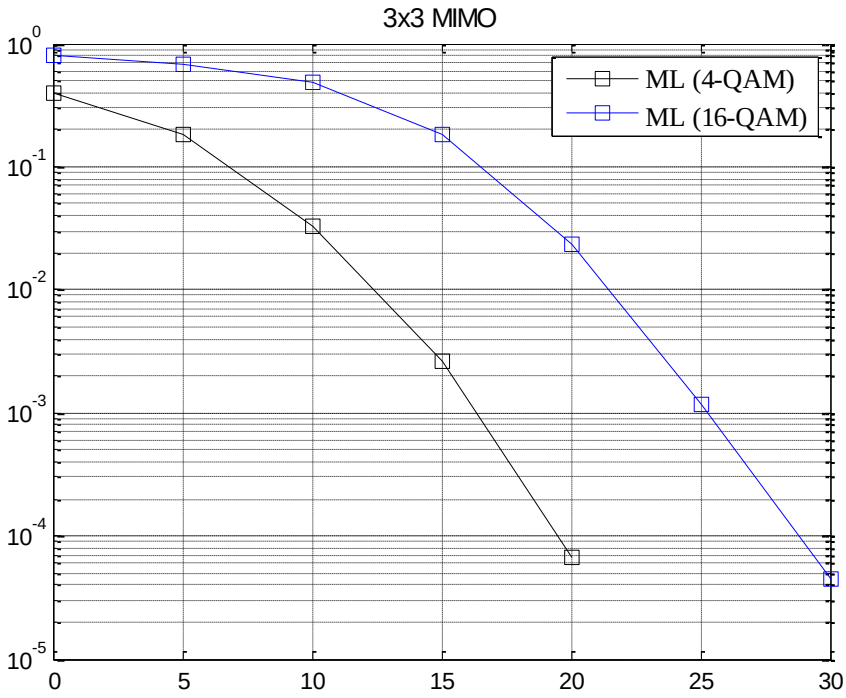

Figure 16 – ML detection for 3 x 3 antennas.

Section 3.4 – Linear Equalization detection

Linear receivers consist of applying a linear transformation to the received vector followed by a quantization to the symbol alphabet (slicing). A simple way to obtain an estimate $\hat{\mathbf{x}}$ is to form

$$\mathbf{x}_{\mathrm{ZF}} = \mathbf{H}^{-1}\mathbf{y} = \mathbf{H}^{-1}\mathbf{H}\mathbf{x} + \mathbf{H}^{-1}\mathbf{n} = \mathbf{x} + \mathbf{H}^{-1}\mathbf{n} \tag{18}$$

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = \mathrm{Q_C}[\mathbf{x}_{\mathrm{ZF}}] \tag{19}$$

This is the so-called linear *zero-forcing* (ZF) receiver, since the interference caused by H is forced to be zero. $\mathbf{H}^{\dagger}$, also known as Moore-Penrose matrix, corresponds to the pseudo-inverse matrix in the complex domain.

$$\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}} \tag{20}$$

The superscript $(.)^{\mathrm{H}}$ denotes Hermitian operator (conjugation followed by transposition or vice-versa). The inversion of $\mathbf{H}$ is a trivial operation however it can only be

defined for invertible matrices, i.e. matrices with non-zero determinant. Here resides the need for $N_R \geq N_T$. The filtered noise is transformed by $\mathbf{H}^{\dagger}$, which constitutes a noise enhancement factor.

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = \mathrm{Q_C}[\mathbf{H}^{\dagger}\mathbf{y}] = \mathrm{Q_C}[\mathbf{H}^{\dagger}\mathbf{x} + \mathbf{H}^{\dagger}\mathbf{n}] = \mathrm{Q_C}[\mathbf{x} + \mathbf{H}^{\dagger}\mathbf{n}] \tag{21}$$

Summarizing the ZF receiver structure has a linear transformation $\mathbf{W}_{\mathrm{ZF}} = \mathbf{H}^{\dagger}$ followed by a quantization to the symbol alphabet by threshold decision

$$\hat{\mathbf{x}}_{\mathrm{ZF}} = \arg\min_{\mathrm{x}\in\mathrm{C^N}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{22}$$

Notice that in $\hat{\mathbf{x}}_{\mathrm{ZF}}$ the search is now made in the continuous domain C^N instead of the discrete complex alphabet C as it is in $\hat{\mathbf{x}}_{\mathrm{ML}}$. This is the origin of the sub-optimality of the ZF receiver. After the inverse transformation, all the points in the lattice are matched back to the initial$\mathbb{Z}^2$. The orthogonal geometry of $\mathbb{Z}^2$ eliminates all the interference between dimensions of the lattice, i.e., between the MIMO layers.

ZF solves the CVP by relaxing it to a search in a continuous neighborhood instead of computing the distance between the target (received vector) and every point in the lattice. In a geometrical perspective ZF performs a linear transformation of the Voronoi regions (mentioned in 2.4) of the $\mathbb{Z}^2$original cubic lattice, by **H**. The resulting regions are called ZF decision regions and correspond to the space where a lattice point will be interpreted as being closer to the lattice point associated with that region [27].

Different bases will output different ZF decision regions that could correspond better or worse to the ideal Voronoi regions of the lattice. The ZF decisions regions are the fundamental regions of the lattice that were shown in section 2.3 and the lower the match between ZF and Voronoi decision regions the greater the SER obtained. Taking into account these facts it is possible to conclude that finding bases such that form the same lattice and better approximate their fundamental regions to the Voronoi decision regions will improve the performance of linear receivers. This is the main idea of the LRA receivers that will be discussed latter.

The noise enhancement factor can be minimized with a minimum mean-square error (MMSE) receiver. In this receiver it is taken into account both the interference and the

noise in order to minimize the expected error. The MMSE linear receiver looks for a linear transformation $\mathbf{W}_{\mathrm{MMSE}}$ that minimizes the mean square error between the estimated vector and the original vector,

$$\mathbf{W}_{\mathrm{MMSE}} = \arg\min_{\mathbf{W}} \mathrm{E}\{\|\mathbf{W}\mathbf{y} - \mathbf{x}\|^2\} \tag{23}$$

There are other possible expressions for representing $\mathrm{W}_{\mathrm{MMSE}}$ but the one assumed in this work is

$$\mathbf{W}_{\mathrm{MMSE}} = \sigma_{\mathrm{x}}^2 \mathbf{I}_{\mathrm{N}} \cdot \mathbf{H}^{\mathrm{H}} \left(\sigma_{\mathrm{n}}^2 \mathbf{I}_{\mathrm{N}} + \mathbf{H} \cdot \sigma_{\mathrm{x}}^2 \cdot \mathbf{I}_{\mathrm{N}} \cdot \mathbf{H}^{\mathrm{H}}\right)^{-1} = \mathbf{H}^{\mathrm{H}} \left\{ \mathbf{H}\mathbf{H}^{\mathrm{H}} + \frac{\sigma_{\mathrm{n}}^2}{\sigma_{\mathrm{x}}^2} \mathbf{I}_{\mathrm{N}} \right\} \tag{24}$$

The estimated $\hat{\mathbf{x}}_{\mathrm{MMSE}}$ is obtained by a linear transformation $\mathbf{W}_{\mathrm{MMSE}}$ followed by a quantization step that is the same as the one used in the ZF case.

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = \mathrm{Q_C}[\mathbf{W}_{\mathrm{MMSE}}\, \mathbf{y}] = \mathrm{Q_C}[\mathbf{W}_{\mathrm{MMSE}}\, \mathbf{H}\mathbf{x} + \mathbf{W}_{\mathrm{MMSE}}\mathbf{n}] = \mathrm{Q_C}[\mathbf{x} + \mathbf{W}_{\mathrm{MMSE}}\mathbf{n}] \tag{25}$$

The MMSE performs better than ZF receiver because it solves the CVP problem by relaxing the search in the continuous space where $\mathbf{x} \in \mathrm{C}^{\mathrm{N}}$ but also for introducing a term that penalizes large $\|\mathbf{x}\|$ and is proportional to the energy of the noise [27].

In order to make an easier implementation an extended system was assumed [12].

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \frac{\sigma_{\mathrm{n}}}{\sigma_{\mathrm{x}}} \mathbf{I}_{\mathrm{N}} \end{bmatrix}, \qquad \bar{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ 0_{\mathrm{N}\times 1} \end{bmatrix} \tag{26}$$

Therefore the estimated $\hat{\mathbf{x}}_{\mathrm{MMSE}}$ can be easily obtained as

$$\hat{\mathbf{x}}_{\mathrm{MMSE}} = \mathrm{Q_C}[\,\bar{\mathbf{H}}^{\dagger}\bar{\mathbf{y}}]\,. \tag{27}$$

Figures Figure 17 to Figure 20 show the performance of ZF and MMSE linear receivers in different configurations. The ML performance is also present as term of comparison with the best performance attainable. Analysing the figures it possible observe that although MMSE receiver performs better than the ZF both SER curves settle to a slope of −1 as proven in literature, which corresponds to a decrease of symbol error rate by a factor of 10 for a 10-fold increase of SNR. From Figures Figure 17 to Figure 20 one can observe that MMSE detection has a *power gain* of some dB compared with ZF, i.e., it

achieves the same SER value from the 4-QAM constellation in a lower value of SNR, yet for higher constellations this power gain decreases as well as for higher SNR values, becoming minimal.

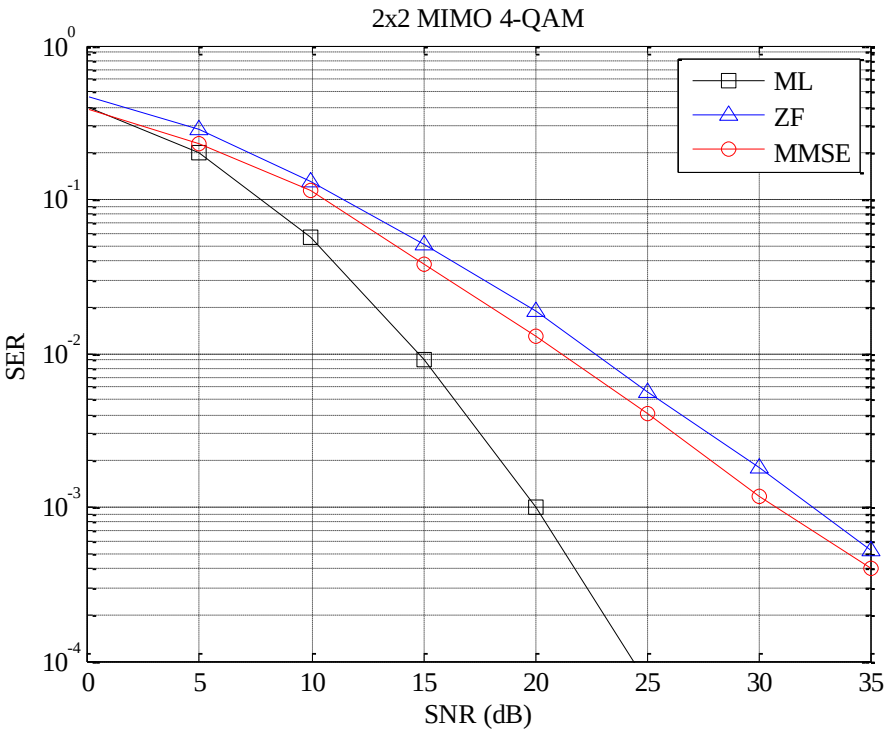

Figure 17 – 4-QAM detections for 2 x 2 antennas using linear receivers.

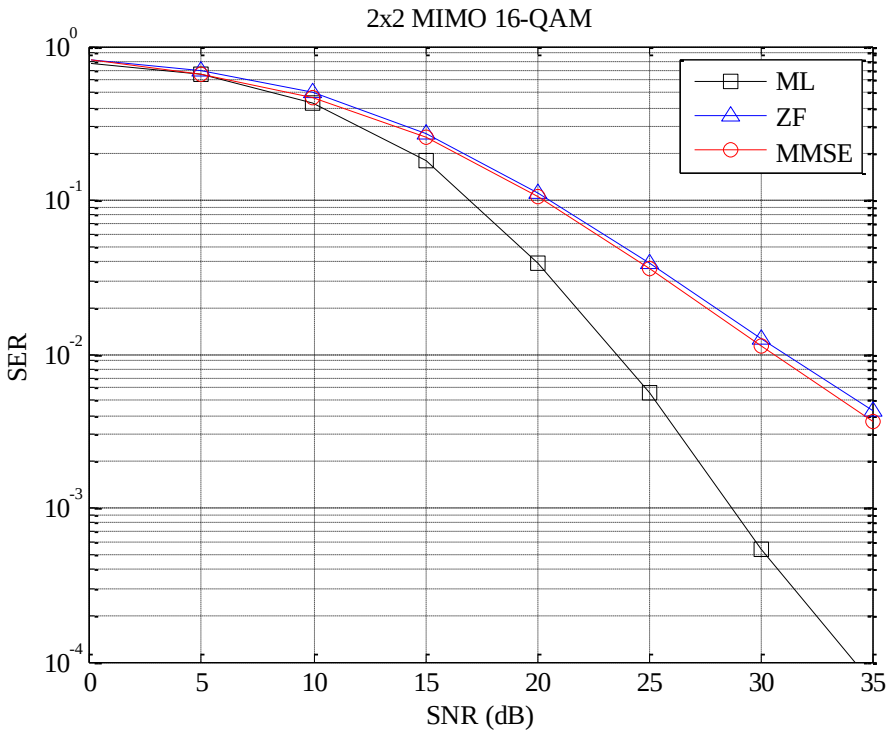

Figure 18 – 16-QAM detections for 2 x 2 antennas using linear receivers.

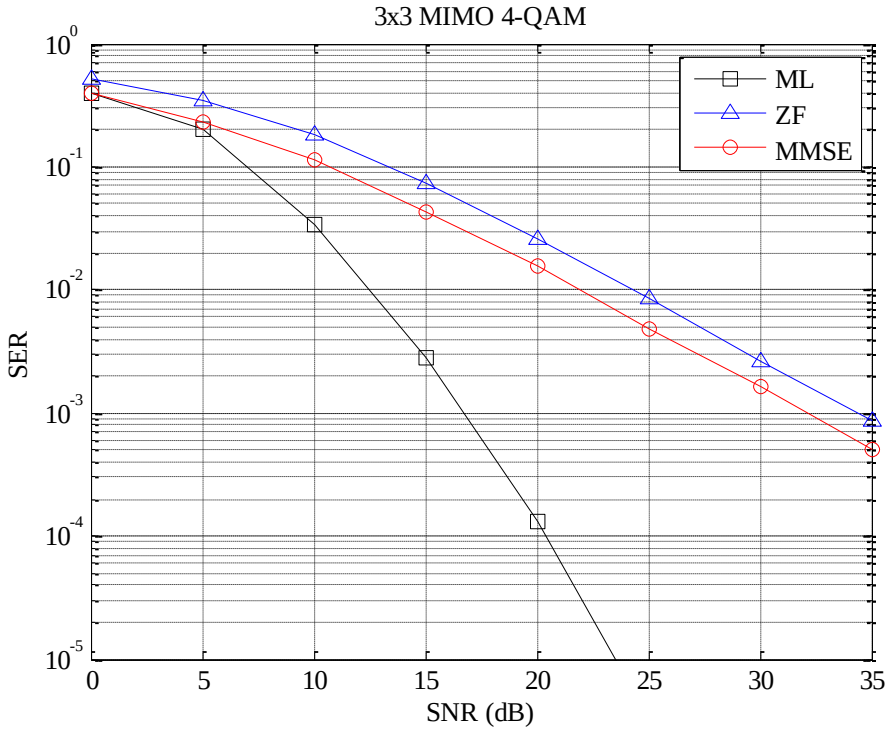

Figure 19 – 4-QAM detections for 3 x 3 antennas using linear receivers

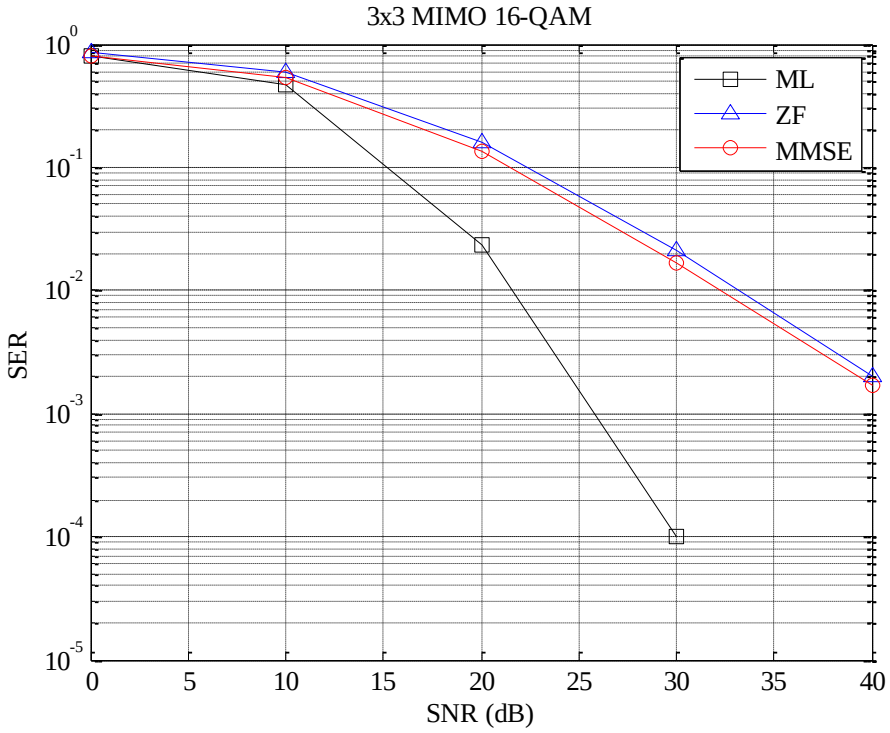

Figure 20 – 16-QAM detections for 3 x 3 antennas using linear receivers.

Section 3.5 – Order Successive interference cancelation detection

Nonlinear methods on the receivers have to involve decision feedback somehow in order to remove the interference between the signals from each antenna. In this context the most commonly used approach is called order successive interference cancelation (OSIC). OSIC in fact corresponds to the original algorithm proposed for the detection of the original V-BLAST transmission scheme in [28]. As it was described in 3.1, V-BLAST allows for the transmission of independent data streams over the matrix channel **H**.

While the transmission side of spatial multiplexing V-BLAST is straightforward, the detection side is more difficult to implement since the signals from each antenna are received by all the receive antennas. The principle of OSIC is to use linear detection to detect first the modulation symbol of the layer least affected by noise and, subsequently, assuming that the first symbol was correctly detected, the interference created by that

symbol is replicated and subtracted from the others layers. The procedure continues detecting the next best signal (again in the sense of the one with the least noise enhancement) and subtracting the interference caused by it to the remaining layers, and repeating this procedure until all symbols are detected.

The optimal criterion at each stage is to select the layer that less enhances the noise power after the linear detection. It was shown in [27] that this corresponds to deciding for the layer that spans the *n*-dimensional lattice by translating parallel hyperplanes of *n*-1 dimensions which have the maximum separation among all possible families of hyperplanes. In doing this, the error probability is minimized for that layer. This corresponds to ordered SIC, without which the performance of SIC is seriously degraded by 3dB [29].

Since the worst subchannel dominates the average SER this is the optimal criterion if all subchannels use the same signaling rate and power. This result was already found in [28], in the first implementation of the V-BLAST detector.

For a better understanding of the OSIC procedure let us follow a practical example of its application on this work for 2×2 MIMO system through 4-QAM constellation and using ZF detection.

$$\mathbf{y}^{(1)} = \begin{bmatrix} -2.43 + 0.08\mathrm{i} \\ +1.07 - 0.97\mathrm{i} \\ -0.69 - 0.06\mathrm{i} \end{bmatrix} = \underbrace{\begin{bmatrix} -0.73 + 0.45\mathrm{i} & -0.10 - 0.74\mathrm{i} & 0.73 - 0.29\mathrm{i} \\ 0.94 - 0.15\mathrm{i} & 0.65 - 0.19\mathrm{i} & -0.24 + 0.70\mathrm{i} \\ -0.30 - 0.61\mathrm{i} & -0.21 - 0.31\mathrm{i} & 0.72 - 0.21\mathrm{i} \end{bmatrix}}_{\mathbf{H}} \mathbf{x} + \mathbf{n}\,. \quad (28)$$

By applying the ZF linear transformations it is obtained

$$\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{y} = \mathbf{x} + \mathbf{H}^{\dagger}\mathbf{n}$$

$$\mathbf{W}^{(1)} = \mathbf{H}^{\dagger} = \begin{bmatrix} -0.18 - 0.58\mathrm{i} & 0.65 + 0.53\mathrm{i} & 1.04 + 0.38\mathrm{i} \\ 0.28 + 0.65\mathrm{i} & 1.59 + 1.28\mathrm{i} & 1.58 - 1.18\mathrm{i} \\ 0.23 - 0.01\mathrm{i} & -0.75 + 1.62\mathrm{i} & 1.92 + 1.94\mathrm{i} \end{bmatrix}. \quad (29)$$

The component of $\hat{\mathrm{x}}$ that experience the lowest noise enhancement by $\mathbf{H}^{\dagger}$ can be detected with high reliability. Since ZMSW noise model is assumed this component corresponds to the row of $\mathbf{H}^{\dagger}$ of smallest *length*, which is the Euclidean norm of the row.

$$\|\mathbf{w}_1(1,:)\|^2 = 2.30 \quad \|\mathbf{w}_1(2,:)\|^2 = 8.55 \quad \|\mathbf{w}_1(3,:)\|^2 = 10.69 \tag{30}$$

The next step is decide the symbol by quantization

$$\hat{\mathbf{x}}(1) = \mathrm{Q_C}\left[\mathbf{W}^{(1)}\mathbf{y}^{(1)}\right] = 1 + 1\mathrm{i}\,. \tag{31}$$

Now the interference due to $\hat{\mathbf{x}}(1)$ is replicated and then subtracted on the others layers

$$\mathbf{y}^{(2)} = \mathbf{y} - \mathbf{H}(:,1)\hat{\mathbf{x}}(1) = \begin{bmatrix} -1.26 + 0.37\mathrm{i} \\ -0.02 - 1.76\mathrm{i} \\ -1.01 - 0.85\mathrm{i} \end{bmatrix} \tag{32}$$

Once the symbol $\hat{\mathbf{x}}(1)$ is already decided the first generator vector in the channel matrix $\mathbf{H}$ becomes zero. Thereby the matrix $\mathbf{H}$ took into account to the second round of the algorithm is (33) and its associated ZF linear transformation matrix (34).

$$\mathbf{H}^{(2)} = \begin{bmatrix} 0 & -0.10 - 0.74\mathrm{i} & 0.73 - 0.29\mathrm{i} \\ 0 & 0.64 - 0.19\mathrm{i} & -0.24 + 0.70\mathrm{i} \\ 0 & -0.21 - 0.31\mathrm{i} & 0.72 - 0.21\mathrm{i} \end{bmatrix} \tag{33}$$

$$\mathbf{W}^{(2)} = \begin{bmatrix} 0 & 0 & 0 \\ 0.92 + 1.14\mathrm{i} & 0.47 + 1.21\mathrm{i} & 0.16 - 0.78\mathrm{i} \\ -0.27 + 0.99\mathrm{i} & -0.94 + 0.09\mathrm{i} & 1.00 + 0.13\mathrm{i} \end{bmatrix} \tag{34}$$

Follows the calculation of the component of $\hat{\mathbf{x}}$ that experience the lowest noise but this time over the estimates available via $\mathbf{W}^{(2)}$ linear equalization of $\mathbf{H}^{(2)}$.

$$\|\mathbf{w}_2(1,:)\|^2 = 0 \quad \|\mathbf{w}_2(2,:)\|^2 = 4.46 \quad \|\mathbf{w}_2(3,:)\|^2 = 2.96 \tag{35}$$

At this stage the third row is the one that less enhances the noise and therefore $\hat{\mathbf{x}}(3)$ is now decided as

$$\hat{\mathbf{x}}(3) = \mathrm{Q_C}\left[\mathbf{W}^{(2)}\mathbf{y}^{(2)}\right] = -1 + 1\mathrm{i}\,, \tag{36}$$

and its interference is subtracted from $\mathbf{y}^{(2)}$

$$\mathbf{y}^{(3)} = \mathbf{y}^{(2)} - \mathbf{H}(:,3)\hat{\mathbf{x}}(3) = \begin{bmatrix} -0.82 - 0.65\mathrm{i} \\ 0.44 - 0.82\mathrm{i} \\ -0.51 - 0.08\mathrm{i} \end{bmatrix} \tag{37}$$

and the matrix **H** becomes in (38), where its third column is now null.

$$\mathbf{H}^{(3)} = \begin{bmatrix} 0 & -0.10 - 0.74\mathrm{i} & 0 \\ 0 & +0.64 - 0.19\mathrm{i} & 0 \\ 0 & -0.21 - 0.31\mathrm{i} & 0 \end{bmatrix} \tag{38}$$

By applying the ZF linear transformations over $\mathbf{H}^{(3)}$ it is obtained $\mathbf{W}^{(3)}$.

$$\mathbf{W}^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ -0.09 + 0.65\mathrm{i} & 0.56 + 0.17\mathrm{i} & -0.19 - 0.27\mathrm{i} \\ 0 & 0 & 0 \end{bmatrix} \tag{39}$$

Finally remains estimate $\hat{\mathbf{x}}(2)$ and the process comes to an end.

$$\hat{\mathbf{x}}(2) = \mathrm{Q_C}\left[\mathbf{W}^{(3)}\mathbf{y}^{(3)}\right] = -1 - 1\mathrm{i}\,. \tag{40}$$

As expected all the symbols were correctly detected. However, one of the basic problems of this strategy is that if one symbol is incorrectly detected, and if it happens to bethe first one to be decided, it leads to a higher probability of incorrect detections of the following symbols. Instead of using a ZF equalization it is also possible perform linear MMSE equalization leading to an MMSE version of the OSIC receiver.

The performance of OSIC receivers in different configurations can be found in Figure 21 to Figure 24. In these figures are present both versions of OSIC receivers, the OSIC combined with ZF (OSIC-ZF) and with MMSE (OSIC-MMSE) for 2×2 and 3×3 MIMO systems using 4 and 16 QAM constellation. Moreover they can be compared with the ML detection. Although the performance of the OSIC receiver capture the same diversity of ordinary linear receivers, $d = 1$, by inspection of Figure 17 to Figure 20 one can conclude that it provides a large power gain when compared with the ordinary linear receivers. Besides with the use of OSIC the MMSE equalizer performance presents a higher difference of power gain to OSIC-ZF then the simple MMSE compared with the simple ZF. One can conclude that the use of OSIC enhance more the performance of MMSE than ZF equalizer.

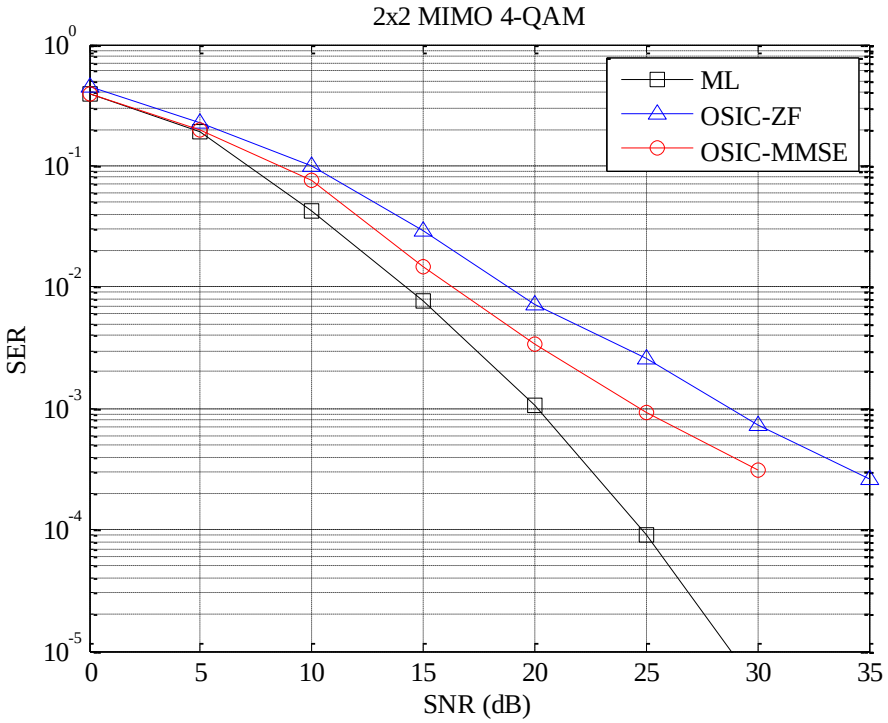

Figure 21 – 4-QAM detections for 2 x 2 antennas using OSIC linear receivers.

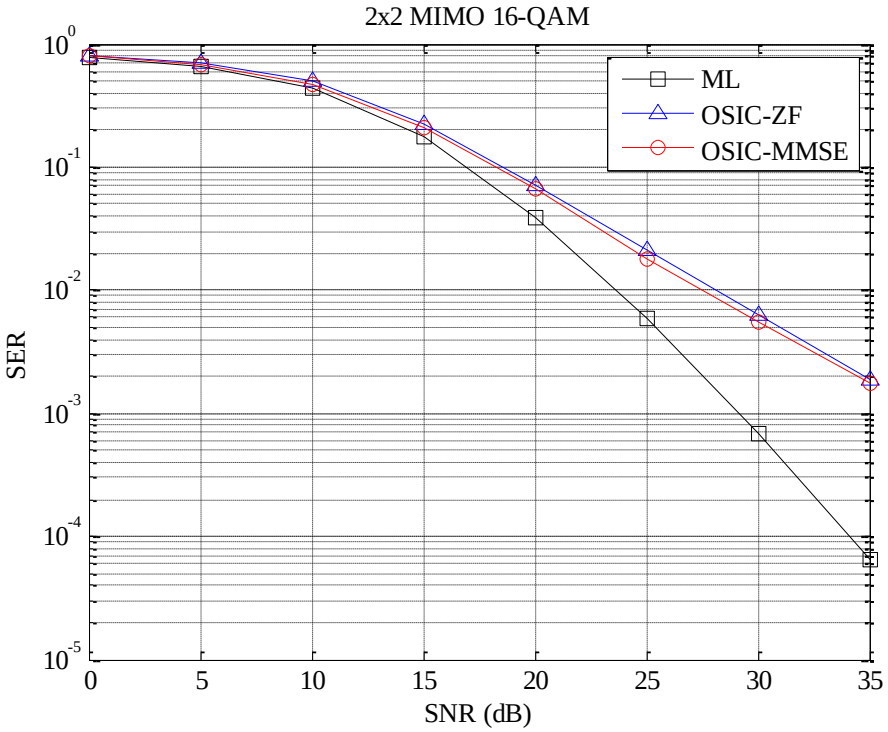

Figure 22 – 16-QAM detections for 2 x 2 antennas using OSIC linear receivers.

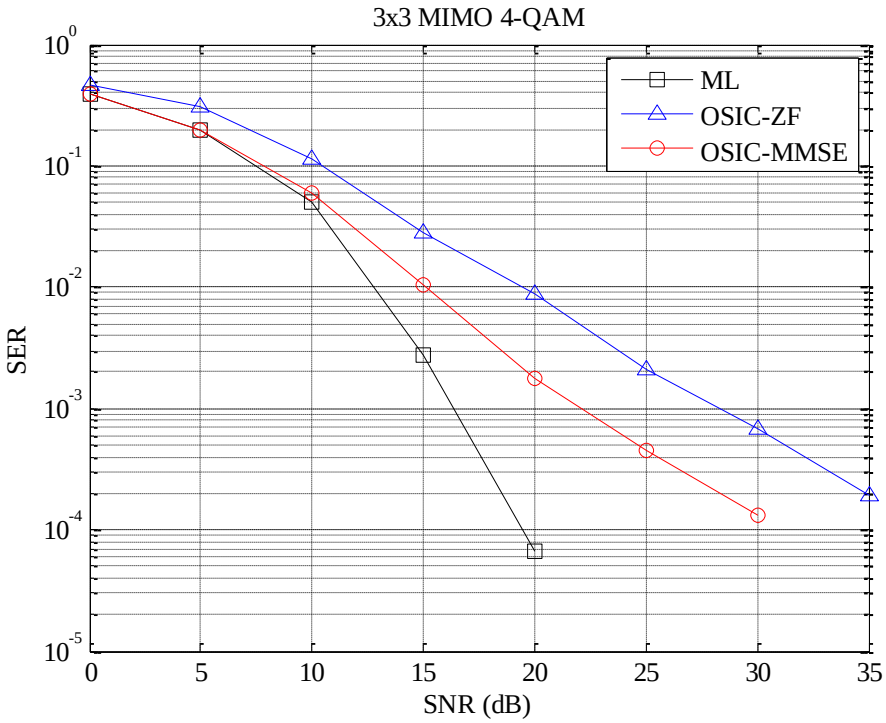

Figure 23 – 4-QAM detections for 3 x 3 antennas using OSIC linear receivers.

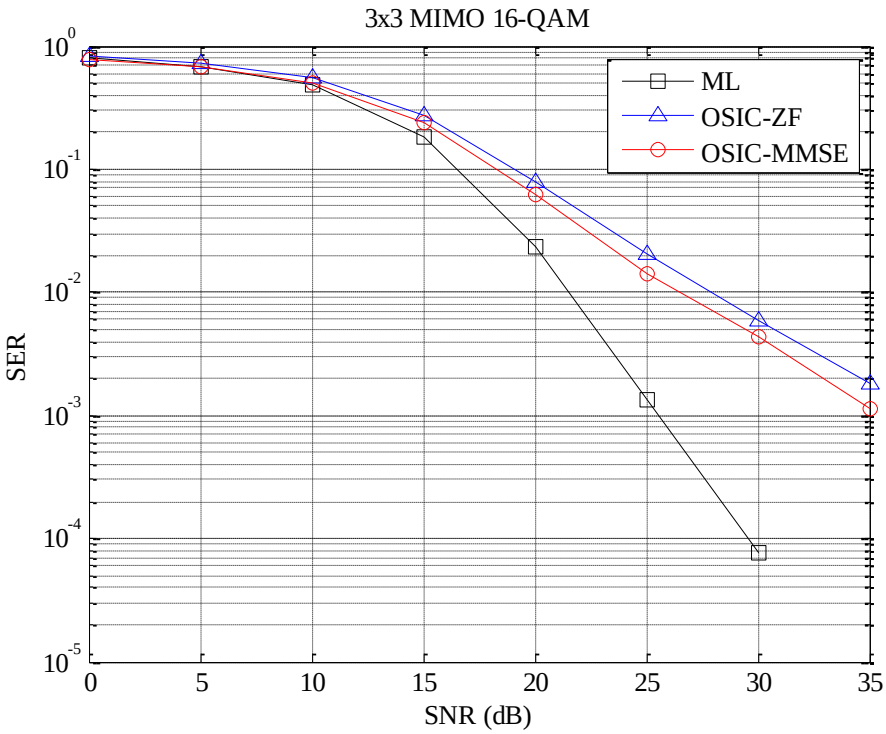

Figure 24 – 16-QAM detections for 3 x 3 antennas using OSIC linear receivers.

Section 3.6 – Lattice reduction-aided

The concept of lattice reduction-aided (LRA) detection arises naturally under this lattice perspective. Considering the detection problem as a CVP, one want s to look for methods that produce good approximate solutions (preferably, in polynomial time). The basic idea behind LRA is to use lattice reduction in conjunction with traditional low-complexity decoders.

Using the lattice detection interpretation, given that the used constellations are QAM constellations, and therefore $\mathbf{x} \in \mathrm{C}^{\mathrm{N}} \subset \mathrm{Z}^{\mathrm{N}}$, the noiseless received points in the communication scenario corresponds to points of the lattice $\mathbf{H}\mathrm{Z}^{\mathrm{N}}$.

As was already discussed different bases can generate the same lattice, however there are bases that perform linear equalizer (ZF and MMSE) decision regions which are closer to the ideal Voronoi regions than others. Lattice reduction can be implemented by algorithms such as LLL and KZ. These algorithms optimise the generating matrix of the lattice to obtain a "better" description of the lattice [21]. The lattice points corresponds to integer linear combinations of the columns of $\mathbf{H} = [\mathrm{h}_1, \ldots, \mathrm{h}_{\mathrm{N}}]$, and lattice reduction obtains another set of column vectors, collected in the reduced matrix $\mathbf{H}_{\mathrm{red}}$, which spans the same set of points, $\mathbf{H}\mathrm{Z}^{\mathrm{N}} \equiv \mathbf{H}_{\mathrm{red}}\mathrm{Z}^{\mathrm{N}}$ while having its generator vectors more orthogonal to each other than the ones in the original basis. Since the matrix $\mathbf{H}_{\mathrm{red}}$ is composed of more orthogonal vectors, the low-complexity detectors will perform better in this lattice because the noise enhancement factor (which for example in linear equalization amounts to $\mathbf{H}_{\mathrm{red}}^{-1}$), is decreased.

The matrices $\mathbf{H}$ and $\mathbf{H}_{\mathrm{red}}$ are related by

$$\mathbf{H}_{\mathrm{red}} \equiv \mathbf{H}\mathbf{M}\,, \tag{41}$$

where $\mathbf{M}$ is a unimodular matrix, i.e., has integer entries and $|\det(\mathbf{M})| = 1$, therefore $\mathbf{M}^{-1}$ is also unimodular and with integer entries.

Using equation (41) in (11) results in

$$\mathbf{y} = \mathbf{H}_{\text{red}} \underbrace{\mathbf{M}^{-1}\mathbf{x}}_{\mathbf{z}} + \mathbf{n}\,, \tag{42}$$

and $\mathbf{z}$ can be interpreted as the noiseless received signal points as the points in the lattice described by $\mathbf{H}_{\text{red}}$. The original data vector $\mathbf{x}$ can be recovered from $\mathbf{z}$ nothing that

$$\mathbf{z} = \mathbf{M}^{-1}\mathbf{x} \implies \mathbf{x} = \mathbf{M}\mathbf{z} \tag{43}$$

The M-QAM constellations used in this dissertation are defined without the origin and have non unitary distance between the symbols. In order to apply lattice reduction algorithms it is necessary to make a translation of the constellation, by creating the modified received vector

$$\mathbf{y}_{\text{red}} = \frac{1}{2}(\mathbf{y} + \mathbf{H}\mathbf{p}) \tag{44}$$

where $\mathbf{p}$ denotes a column vector of N elements all equal to 1.

After the lattice reduction pre-processing stage, any of receivers earlier described in this dissertation may be used concatenated with the lattice reduction block, namely ZF, MMSE and OSIC.

The complex LLL (CLLL) was implemented in the simulations with LRA. Since the CLLL reduces basis defined by complex matrices, other components of MIMO detectors, such as the computation of pseudo-inverse and QR decomposition, can also be operated with complex arithmetic without having to double the dimensions of the equivalent channel matrix. Thus, in some sense, the use of CLLL helps reducing the complexity of other parts of the MIMO detector, not just the reduction algorithm.

Table 2 represents the pseudo-algorithm implemented to simulate the LRA receivers.

Table 2 - Pseudo code for LRA detection

1: Shift and scale the constellation to have zero as a lattice point:

$y_{red} = \frac{1}{2}(y + Hp)$,

$p = \left[\underbrace{1,\ldots,1}_{N}\right]^{T}$.

2: Reduce the lattice basis **H** using CLLL:

$[H_{red}, M] = CLLL(H)$.

3: Apply some detector to the CVP defined by (H_{red}, y_{red})

$z = Detect(H_{red}, y_{red})$.

4: Using z and M estimate the symbol in the original coordinate system

$\hat{x} = 2MQ_z[z] - p$.

Figures Figure 25 to Figure 32 show the performance of LRA receivers in different configurations. These configurations include 2×2 and 3×3 MIMO systems using both 4 and 16 QAM constellations. The LRA receivers capture the same diversity of ML receivers as proved in literature. However they suffer some power penalty. By analyzing Figure 25 to Figure 29 one can conclude that LRA ZF an LRA MMSE in 2×2 MIMO system, taking into account the ML detection for each case, have less power penalty when compared with its use in 3×3 MIMO system. Besides the power penalty difference between LRA ZF and MMSE detection becomes higher in the 3×3 MIMO system, been the latter the one that achieves better performance losing only a few dB from the optimal ML detection in 2×2 MIMO system. In the lattice reduction followed by OSIC strategies, depicted in Figure 30 to Figure 32, it is possible to observe the power gain over the LRA linear receivers is large, especially for the case of 3×3 MIMO system. Furthermore the fact observed in section 3.4 is also noticed, once that after the pre-processing stage of lattice reduction the OSIC-MMSE performs much better than the OSIC-ZF detection.

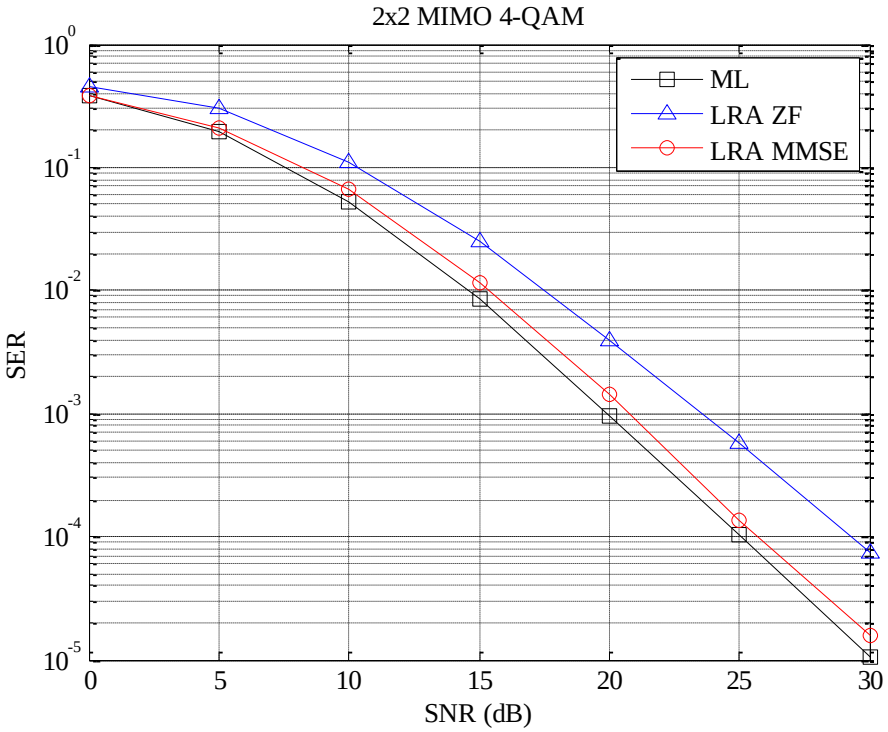

Figure 25 – 4-QAM detection for 2 x 2 antennas using LRA linear receivers.

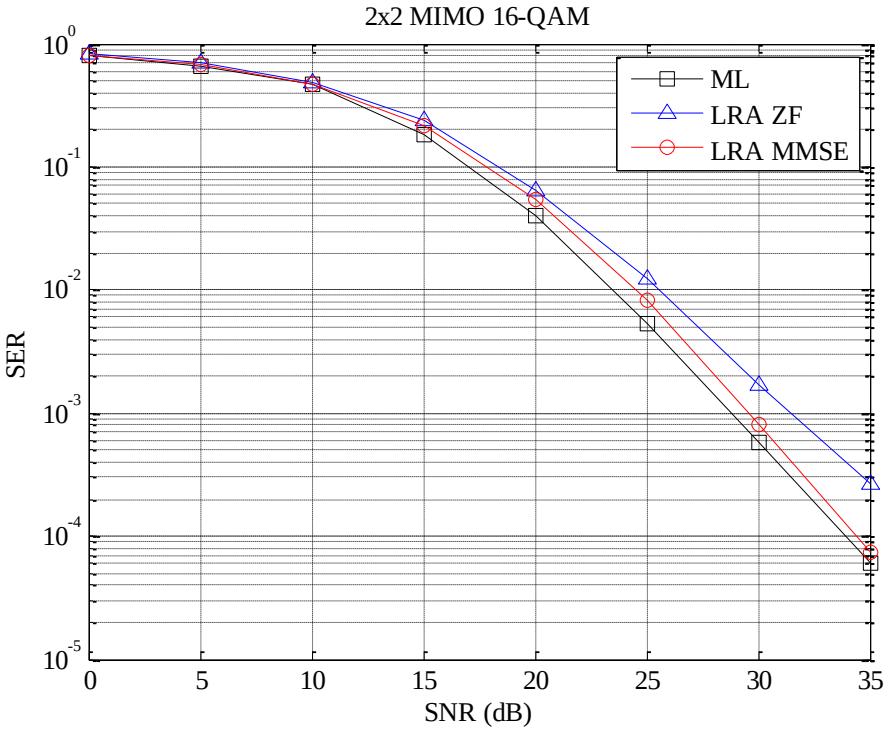

Figure 26 – 16-QAM detection for 2 x 2 antennas using LRA linear receivers.

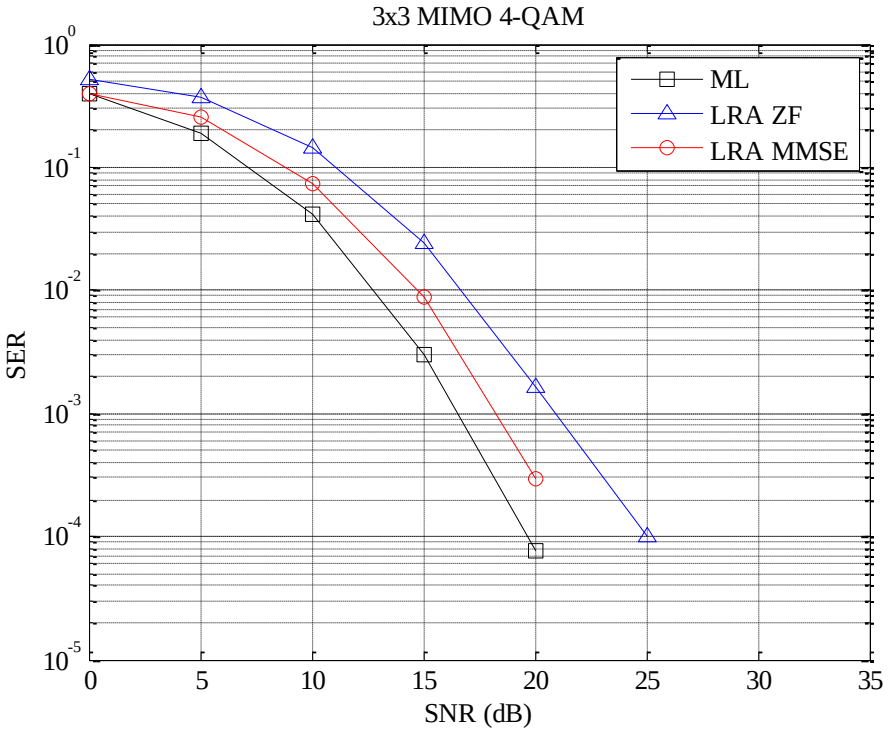

Figure 27 – 4-QAM detection for 3 x 3 antennas using LRA linear receivers.

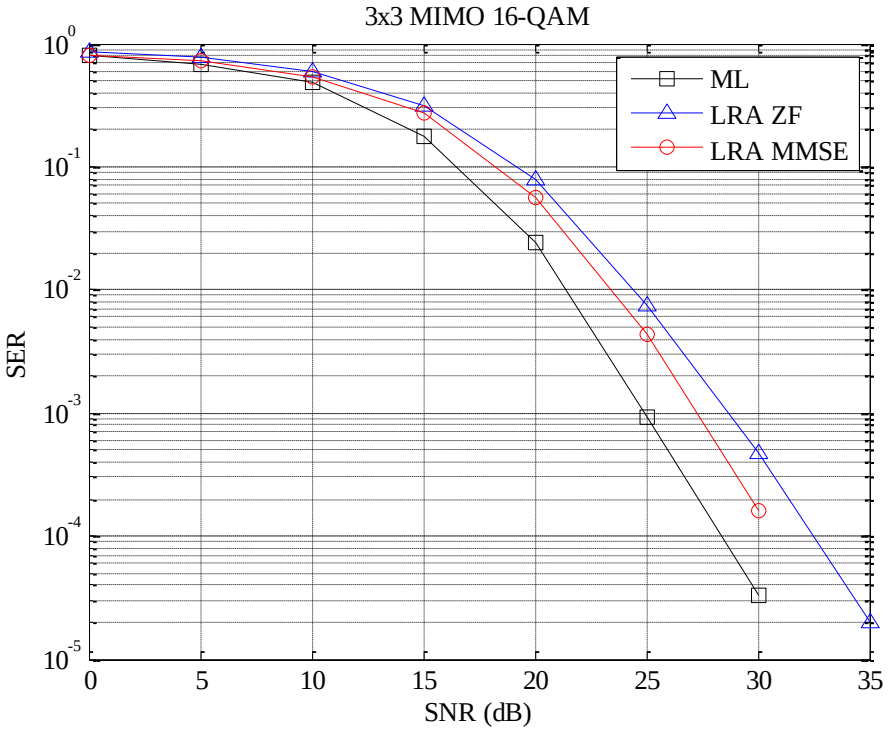

Figure 28 – 16-QAM detection for 3 x 3 antennas using LRA linear receivers.

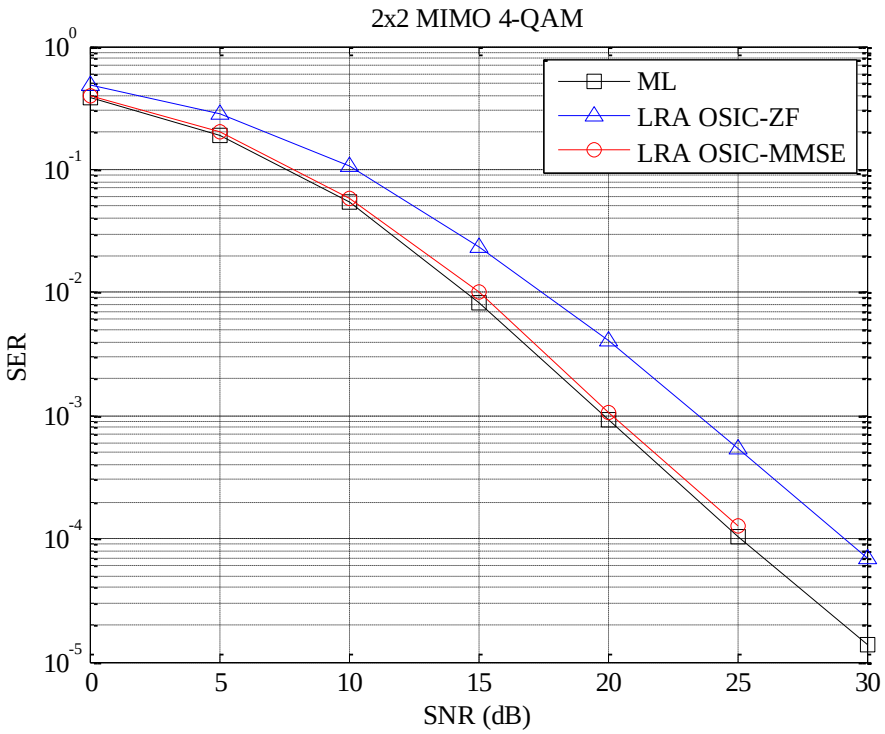

Figure 29 – 4-QAM detection for 2 x 2 antennas using LRA OSIC linear receivers.

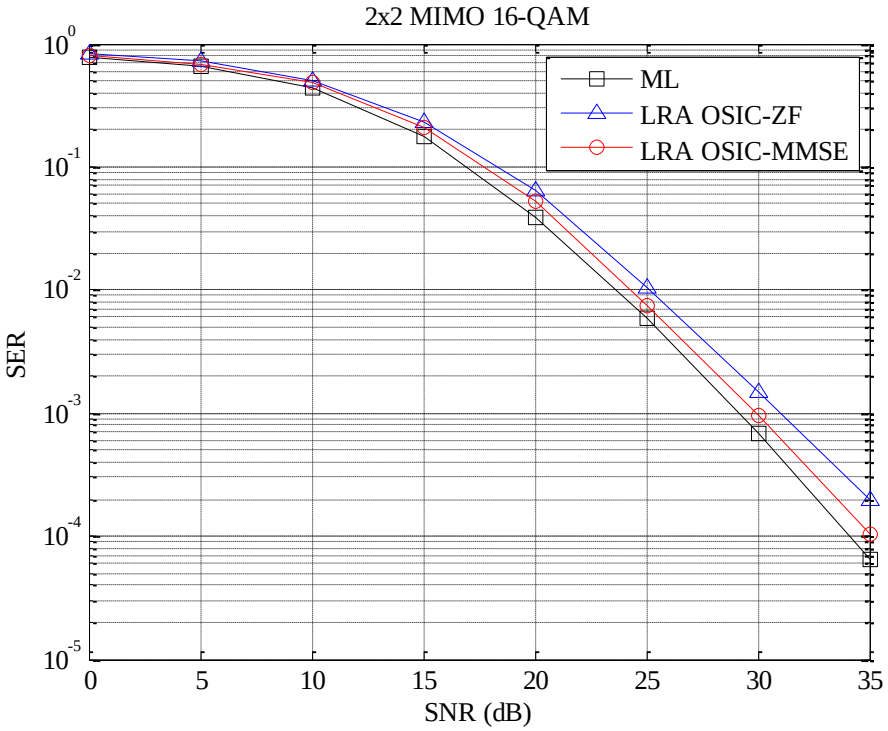

Figure 30 – 16-QAM detection for 2 x 2 antennas using LRA OSIC linear receivers.

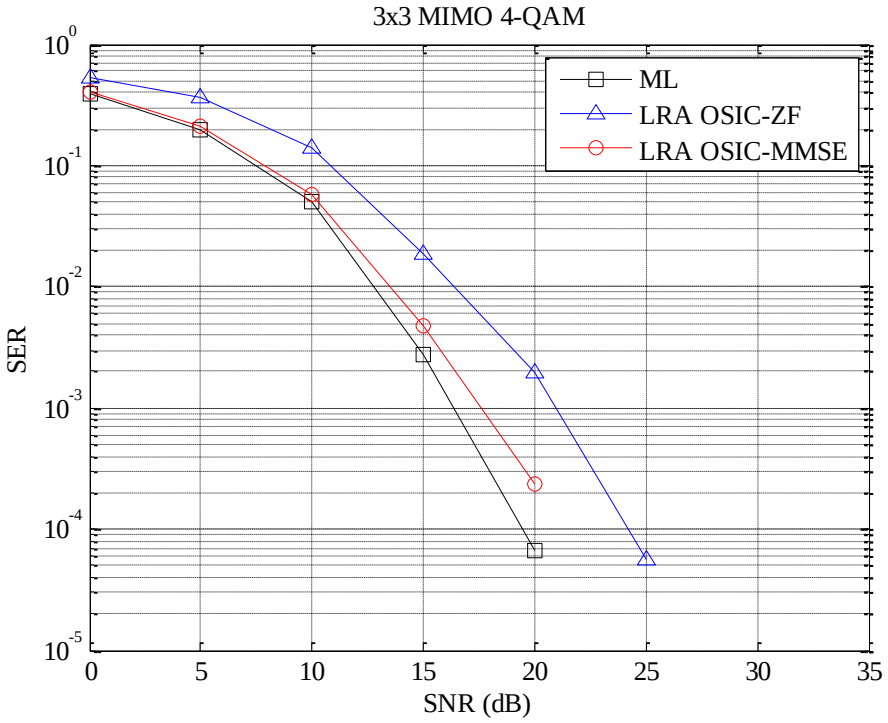

Figure 31 – 4-QAM detection for 3 x 3 antennas using LRA OSIC linear receivers.

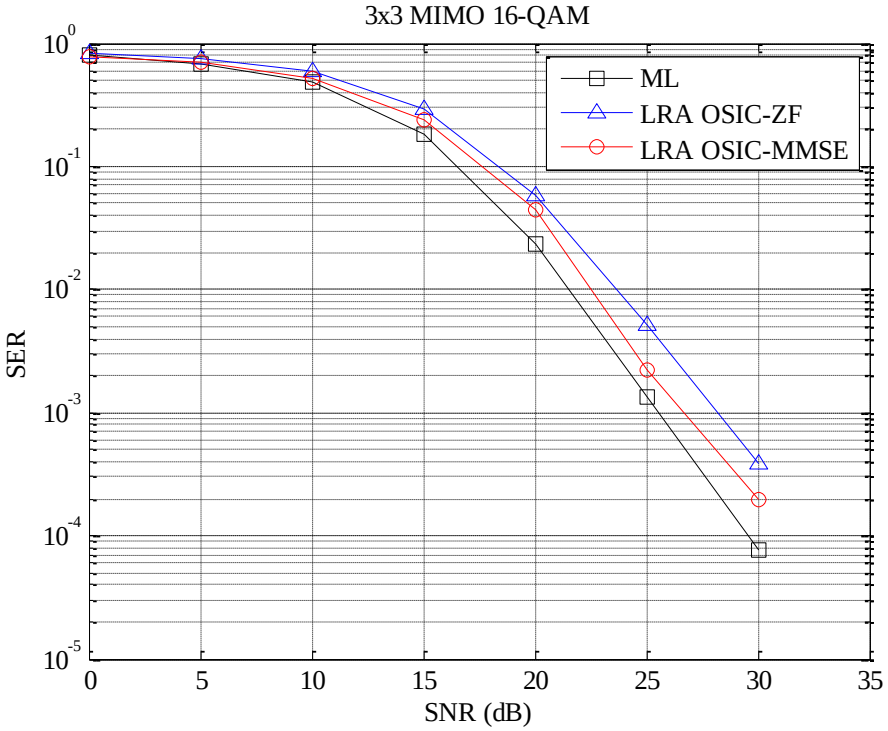

Figure 32 – a 16-QAM detection for 3 x 3 antennas using LRA OSIC linear receivers.

Looking at all the strategies implemented for the MIMO detection so far, the performance results point to the conclusion that LRA OSIC-MMSE receiver is the one that achieves the best performance, as known in the recent literature. It capture the same diversity order of ML detection, $d = N$, and among the other receivers is the one that has the smaller power penalty compared with ML. On the other hand is also the one that needs more computation, so apart from the LRA receivers is OSIC-MMSE receiver whose has the best performance. It has a slope of -1 but is the one that minimizes the power penalty compared with ML detection, again apart from LRA receivers.

Chapter 4

PLNC with MIMO relay

A set of new strategies will be proposed on this Chapter that move beyond the traditional TWRC by merging MIMO and full duplex with PLNC

Section 4.1 – System model

Throughout this chapter physical-layer network coding (PLNC) will be used with MIMO technics, as explain on the previous chapter. In PLNC at the physical layer EM signals can be correctly added with properly modulation and demodulation technics. So understanding the structure of the physical layer is important in this Chapter.

The physical layer considered in this dissertation is the wireless medium and thus its properties will be reviewed, despite the fact that some of them have already been considered implicitly in the previous chapter.

Based on [30] in a PLNC perspective there are three key observations to be made:

- Signal fading is linear, between the transmitter and a receiver, an EM signal experiences a linear transformation. This transformation is primarily caused by reflections and multipath propagation. Assuming band-limited communication the respective signals can be represented uniquely by complex-valued discrete-time samples. Then the received signal at any point in space can be expressed as the convulsion of the transmitted signal with an impulse-response function that characterizes the signal propagation. So a usual approach is to model this impulse response function in a statistical way instead of modelling exactly the physical surroundings.

  Assuming flat (i.e., frequency non-selective) fading, as it would be appropriate for narrow band communications, the channel impulse response function reduces to a delayed Dirac delta function whose height and time-delay characterizes the signal. This time delay may be viewed as a time-shift in the discrete-time simulation model. This particularly simple model is the one considered in the previous chapter and is now applied to a network. Taking into account the concept of time-slot, the induced signal can be expressed as $hx[t]$ where $x[t]$ is the sequence of complex symbols transmitted at the symbol transmission period T_s. The prevalent model in the literature assumes a random fading h taken from a zero-mean circularly

symmetric complex Gaussian distribution with unit variance, as it was also considered in Chapter 3. As well as in the previous chapter the strategies discussed below require CSIR, which is motivated by signal measurements that can be acquired at the receiver. Notice that the transmitter is oblivious to h.

- Multiple signals interfere in a linear additive way. Instead of considering one transmitter only, let us now consider L nodes transmitting simultaneously. The induced signal at any point in space can be expressed as

$$\sum_{l=1}^{L} h_l x_l[t] \,. \tag{45}$$

It is commonly assumed that the respective fading coefficients h_l are independent from each other since each node transmits from a different position in space. Throughout this chapter all those coefficients follow also a zero-mean circularly symmetric complex Gaussian distribution with unit variance and, in order to have an independent and identically distributed Rayleigh fading channel model, the phase of each h is uniformly distributed in $[0\ 2\pi[$, and its amplitude has a Rayleigh distribution.

- Noise is independent of the signal and added at the receiver. At any receiving antenna one has to consider some additive noise, denoted by n which is taken from an independent circularly symmetric complex Gaussian with zero average and variance σ_n^2 corresponding, usually dubbed as the ZMSW noise model (e.g., [Goldsmith 2013]).

Taking into account these three observations along with several other detailed considerations [31] it is possible to define the commonly used model for the signal at a particular receiver when multiple L nodes are simultaneously transmitting as

$$y[t] = \sum_{l=1}^{L} h_l x_l[t] + n[t]. \tag{46}$$

Notice that the errors caused by the noise always accumulate over the various stages of the network, which is an undesirable element for linear network coding at the physical layer.

Perfect synchronization between the nodes is assumed in this dissertation, however, synchronization is for long time an active research problem in wireless networks. Symbol time and carrier-frequency synchronizations, which are both needed in PLNC, have been actively investigated by researchers in fields such as OFDMA [32].

PLNC in TWRC have been intensively researched in the recent PhD dissertation [33], where a strategy exploring the intersection of PLNC, MIMO, and the Alamouti coding scheme are proposed. However, that research did not venture beyond the TWRC. As soon as the theoretical understanding of this scenario matures, it is most likely that the research focus will move towards the application of PLNC to general topologies. Up until now, the literature on this area is scarce and it is difficult to find examples to compare results.

Combine PLNC with MIMO will be started with the TWRC. More specifically, the assumed scenario consists of a certain number of terminals that want to exchange messages among them, while knowing that there is no direct link between them. In some literature (mostly in information theory) this is also known as the *Y-channel* [34]. Therefore they have to revert to a communication made through a relay. All the nodes are considered half duplex and all channels are considered reciprocal.

In all the strategies proposed, the uplink (from the terminals to the relay) consists of decentralized transmitters (mobile terminals) and on a first iteration of the communication protocol, a central receiver (MIMO relay), which in the following phase becomes a central transmitter. Ne starts by considering terminals equipped with a single antenna and a relay equipped with a number of antennas equal to the number of terminals, this can be considered as a so-called *distributed MIMO system* (multipoint-to-point). The uplink direction of any multiuser mobile communication system is also an example of a distributed MIMO system. The role of the base station is to recover the individual users signals from its received signal, and because a number of users transmit at the same time in the same band, this received signal is the superposition of all the active users signals [21]. Figure 33 illustrates the scenario considered in the uplink.

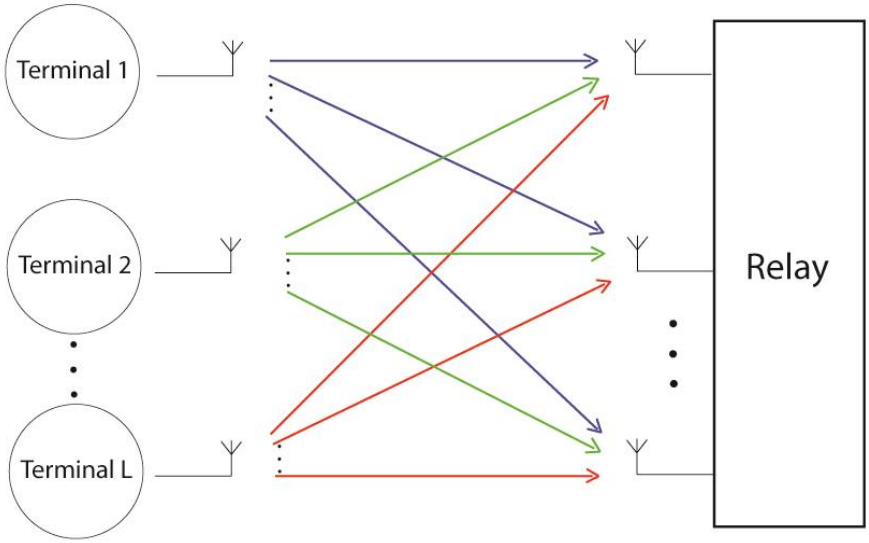

Figure 33 – Distributed MIMO in the uplink.

Since the uplink scenario is a MIMO system, even being a distributed one, all the detection strategies discussed on the previous chapter can be applied, considering perfectly synchronized networks and CSIT.

The MIMO relay receives the sum of the signals coming from the different terminals and detects the message sent by each user. Then, in the downlink, it chooses a strategy to send the messages back to the terminals in order to succeed the exchange of messages between them. After receiving the sum of the signals sent by the MIMO relay each terminal attempts to obtain the messages from the other terminals.

The strategies used at the receivers in the downlink are based in a PLNC perspective since they take advantage of the fact that propagation in wireless medium results in a signal that is a linear combination of the signals sent by the MIMO relay, i.e., the receivers know which are those linear combinations once it is considered CISR and so they can cancel the channel effect. However the decoding results may not be precise due to the induced noise. The signal at each terminal, in the downlink, contains the message that the terminal itself sent in the uplink phase. Consequently each terminal can subtract the contribution of its message from the downlink received signal, what can be seen as self-interference (SI) cancelation. Then each terminal tries to decode the messages from the other ones using this signal.

Besides the traditional TWRC configuration, this dissertation proposes to combine MIMO techniques with PLNC in a scenario with three users and a relay. While in the TWRC case the traditional scheme takes 4 time-slots to accomplish the exchange of messages, as it was explained in Chapter 1, it is straightforward to intuit that in the new scenario, with three terminals and a relay, the exchange of the messages takes 6 time-slots in the traditional TDMA scheme.

There will be considered two strategies for the exchange of messages in the scenario with three terminals: one that uses just 2 time-slots, improving the throughput in 150%, and another one that uses 3 time-slots, which have an improvement of 100% compared to the traditional scheme.

For all the cases discussed the MIMO relay has a high processing capacity and because of that the detection will always use LRA OSIC-MMSE and the uplink takes just one time-slot making use of the distributed MIMO system. This option was taken because LRA OSIC-MMSE was the detection strategy that achieved the best performance among the others in the previous chapter, excluding ML detection that has exponential time consumption.

As it was stated in Chapter 3 the newly implemented strategies performance results will be plotted as *symbol error rate* (SER) as function of the SNR. Notice that for simplicity of results the plots will just have one curve for the uplink, regarding the MIMO relay, and another one for the downlink representing the mean of all terminals.

The SER, both on the uplink and the downlink phase, it is obtained comparing the messages decoded by each terminal with the original messages sent by them in the uplink phase. Notice that the downlink performance suffers from errors due to the previous uplink detection in the relay. This reinforces the choice of LRA OSIC-MMSE in the MIMO relay as an attempt to minimize the impact in downlink performance.

Section 4.2 – TWRC using MIMO combined with PLNC

Figure 34 and Figure 35 is show an illustrated system of MIMO combined with PLNC in TWRC. The two terminals are equipped with a single antenna and the relay with two antennas. The channel coefficients h are also depicted in both Figures and the messages exchanged are denoted by $x \in \mathbb{C}$.

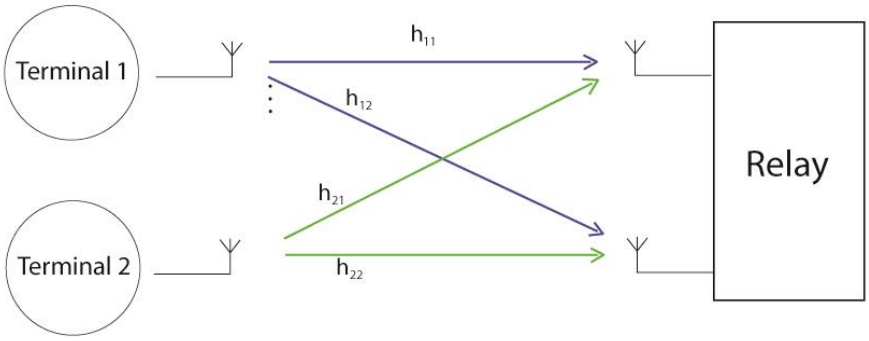

Figure 34 – Uplink in PLNC combined with MIMO inthe TWRC.

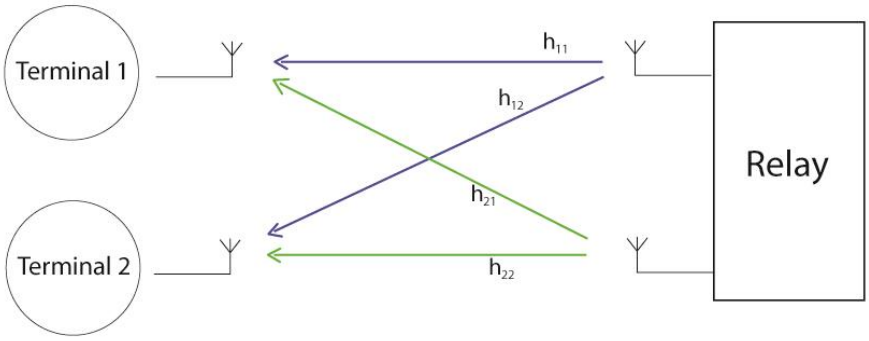

Figure 35 – Downlink in PLNC combined with MIMO in the TWRC.

In the uplink phase terminals 1 and 2 send their messages to the relay. As in an ordinary 2×2 distributed MIMO system the relay receives the messages x_1 and x_2. Then the relay sends x_1 in one his antennas and x_2 in the other. This transmitting configuration is

assumed to be known in advance by the terminals and therefore, accordingly with what was described in section 4.1, the received signal of terminal 1 is given by

$$y_1 = h_{11}x_1 + h_{12}x_2 + n\,. \tag{47}$$

The terminal possesses channel state information (CSIR) and knows its own sent uplink message. Hence, the first step of the decoding process amounts to canceling its own contribution from the received signal, resulting in

$$y_1' = y_1 - h_{11}x_1. \tag{48}$$

Making use of a simple matched filter that cancels the channel linear transformation experienced by message x_2, terminal 1 can obtain a noisy version of the message x_2 and then apply a quantization step,

$$\hat{x}_2 = \mathrm{Q_C}[h_{12}^{-1}\, y_1']\,. \tag{49}$$

The same procedure is followed by terminal 2 to retrieve the desired message. The performance of this strategy is presented hereafter in Figure 36 and Figure 37 for the 4 and 16-QAM constellations, respectively. The performance results of this strategy are poorer than the ones presented in [33] for TWRC, since the Alamouti coding scheme was not used here. However, the throughput of the present strategy is improved by 100% when compared with the traditional scheme discussed in Chapter 1, obtaining a reduction of communication steps from 4 time-slots to 2 time-slots.

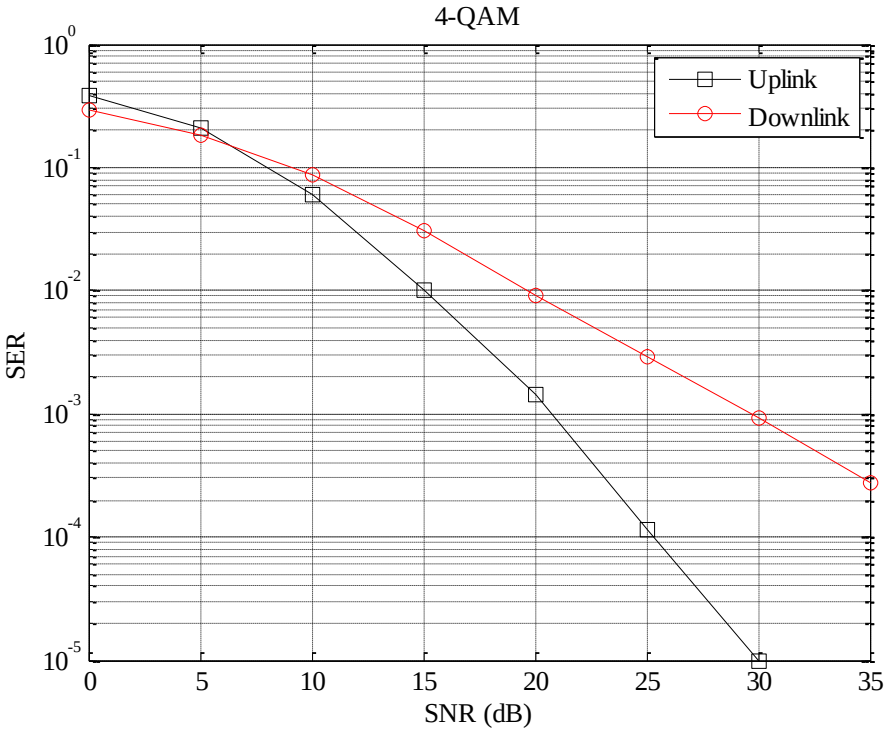

Figure 36 – 4-QAM detection in PLNC combined with MIMO in TWRC.

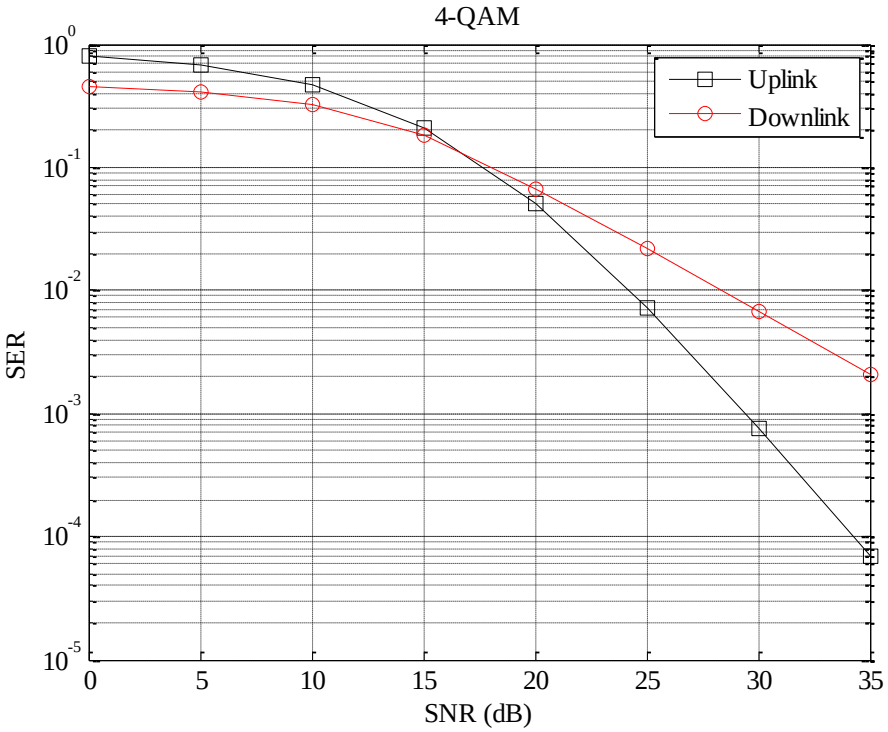

Figure 37 – 16-QAM detection in PLNC combined with MIMO in TWRC.

Section 4.3 – PLNC with MIMO relay, 2time slots strategy

The scenario assumed consists of three terminals which want to exchange messages between them. The messages are symbols $x_i \in C$, where the index i refers to the terminal from where the symbol was originally transmitted in the uplink. As in the previous scheme the terminal cannot hear each other and are only able to transmit do the relay. The uplink consists of a simple case of a distributed 3×3 MIMO system with a relay equipped with three antennas while the terminals only use one antenna to transmit their messages to the relay. This stage takes only one time-slot and the MIMO relay makes use of a LRA OSIC-MMSE receiver to decode the messages sent by the terminals. Figure 38 illustrates the uplink phase and the channel coefficients considered.

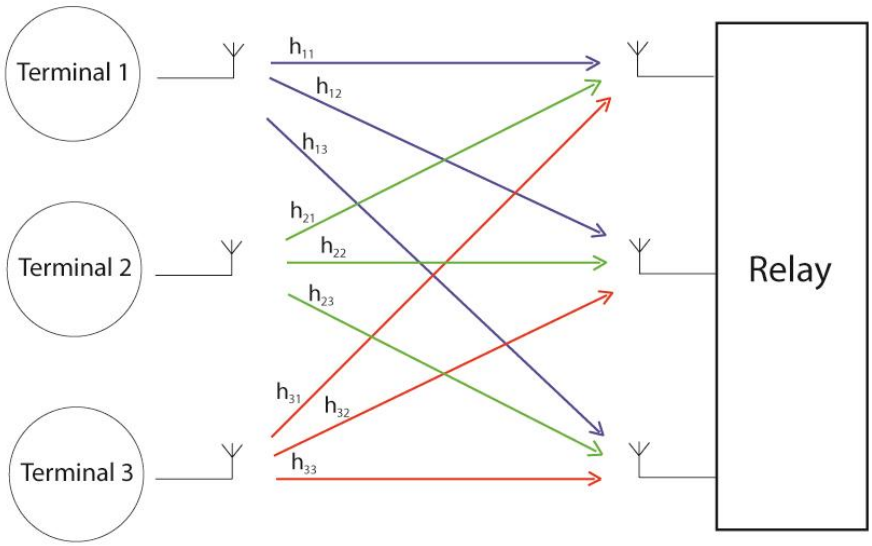

Figure 38 – Uplink with 3 terminals.

In the uplink the MIMO relay decodes the messages x_1, x_2 and x_3 coming from the three terminals. In the two time-slots strategy the downlink phase, shown in Figure 39, consists of the transmission of those messages by antenna 1 of the relay, transmits x_1, antenna 2, transmits x_2 and antenna 3, transmits x_3, all at the same time on the second time-slot.

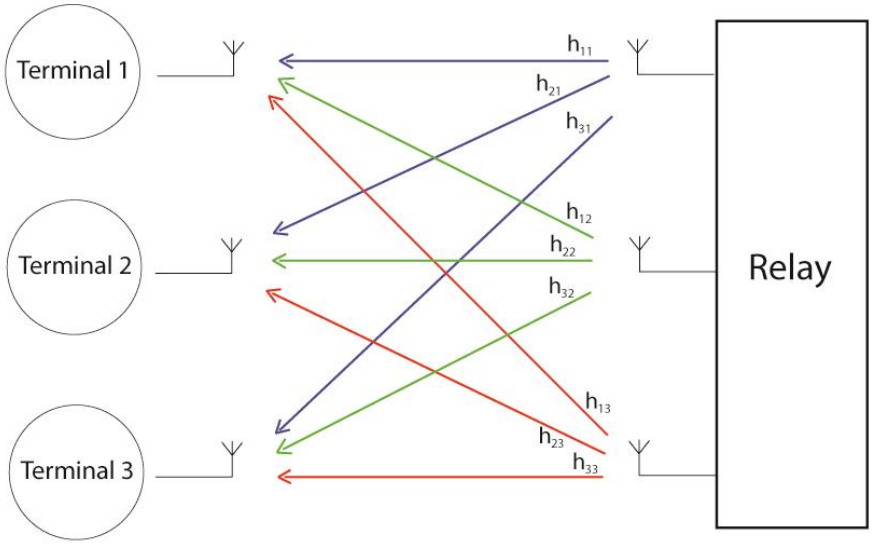

Figure 39 – Downlink for two time-slot strategy.

The transmitting configuration mentioned is assumed to be known in advance by the terminals, after the linear transformation performed by the wireless medium described in section 4.1, the received signal at each terminal is given by

$$y_1 = h_{11}x_1 + h_{12}x_2 + h_{13}x_3 + n\,, \tag{50}$$

$$y_2 = h_{21}x_1 + h_{22}x_2 + h_{23}x_3 + n\,, \tag{51}$$

$$y_3 = h_{31}x_1 + h_{32}x_2 + h_{33}x_3 + n\,. \tag{52}$$

The coefficients h are the ones represented in Figure 39 and y_i denote the received signal at the i^{th} terminal. Since all terminals possess CSIR and knows its own sent uplink message, the first step of the decoding process amounts to canceling its own contribution from the received signal, resulting in

$$y_1' = y_1 - h_{11}x_1\,, \tag{53}$$

$$y_2' = y_2 - h_{22}x_2\,, \tag{54}$$

$$y_3' = y_3 - h_{33}x_3\,. \tag{55}$$

The y_i' obtained are now decoded via joint detection (i.e., brute force or ML), detection allowing each terminal to estimate the other two messages. In this configuration, it is very important to note that this ML detection is made using the projections of the signals in just one dimension, since the y_i' are just a scalar and not vectors. Therefore, signal detection will be much affected by noise given that in this projective space the Euclidean distance between symbols is not lower bounded which eventually contributes in a decisive way to the poor performance of this strategy. It is also worth pointing out that the downlink performance suffers from errors due to the previous uplink detection in the relay.

In Figure 40 and Figure 41 are the performance results for 4 and 16-QAM modulation. The performance of this strategy, apart from the higher complexity due to ML detection, is lower than the one obtained on 4.2, even though it concerns three terminals. Nevertheless, the throughput of the present strategy is improved by 150% compared with the traditional scheme, reducing the communications from 6 time-slots to 2 time-slots.

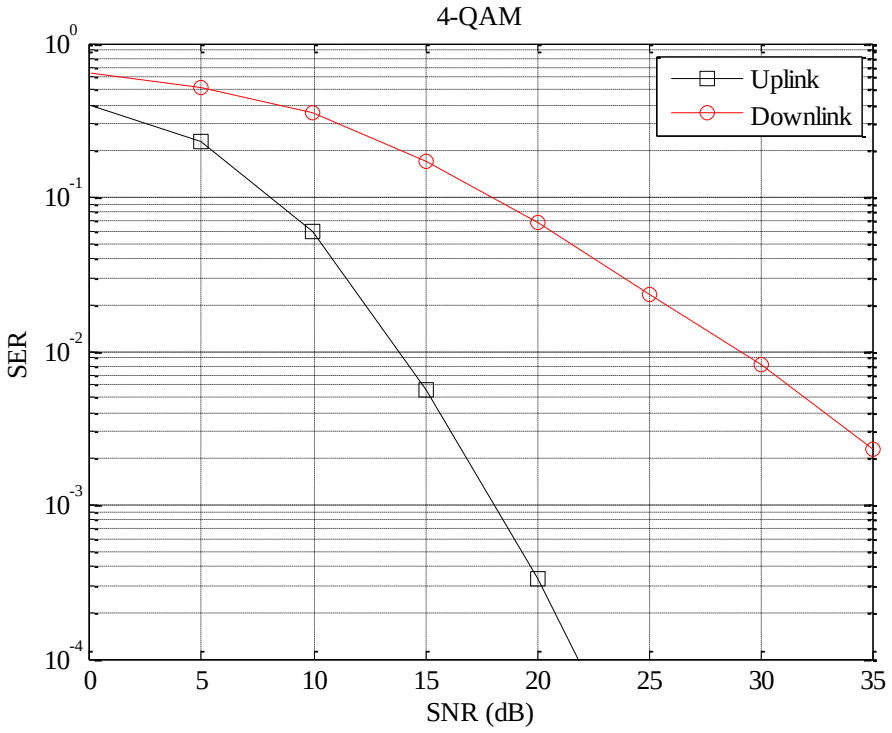

Figure 40 – 4-QAM detection in two time-slot strategy.

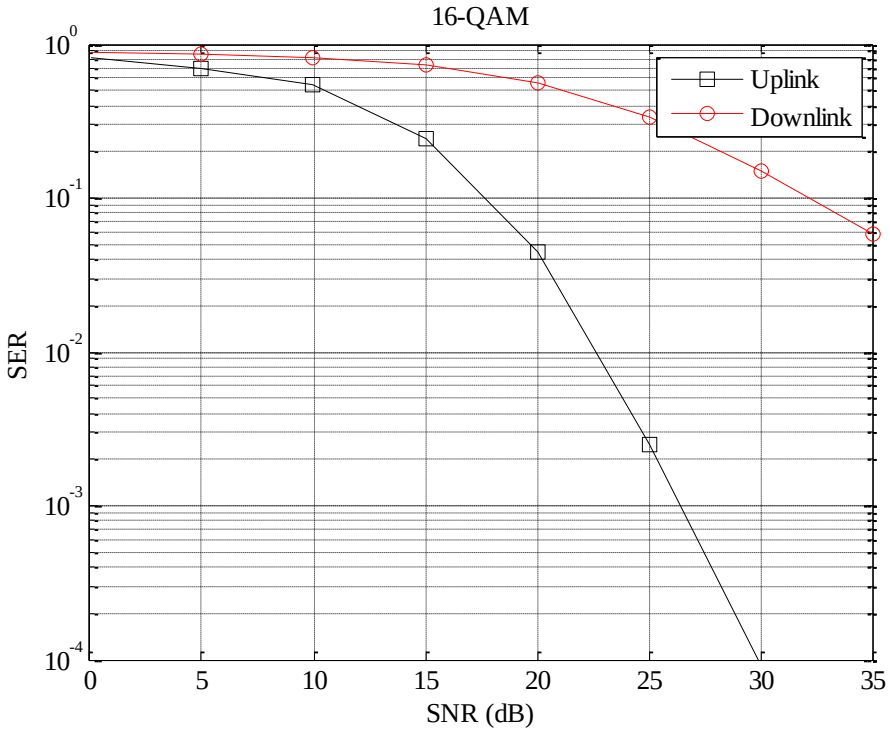

Figure 41 – 16-QAM detection in two time-slot strategy.

Section 4.4 PLNC with MIMO relay, 3time slots strategy

Considering the scenario described on 4.3 for the uplink the MIMO relay decodes the messages x_1, x_2 and x_3 in the first time-slot. This three time-slot strategy uses 2 time-slots in the downlink phase, as illustrated in Figure 42 and Figure 43.

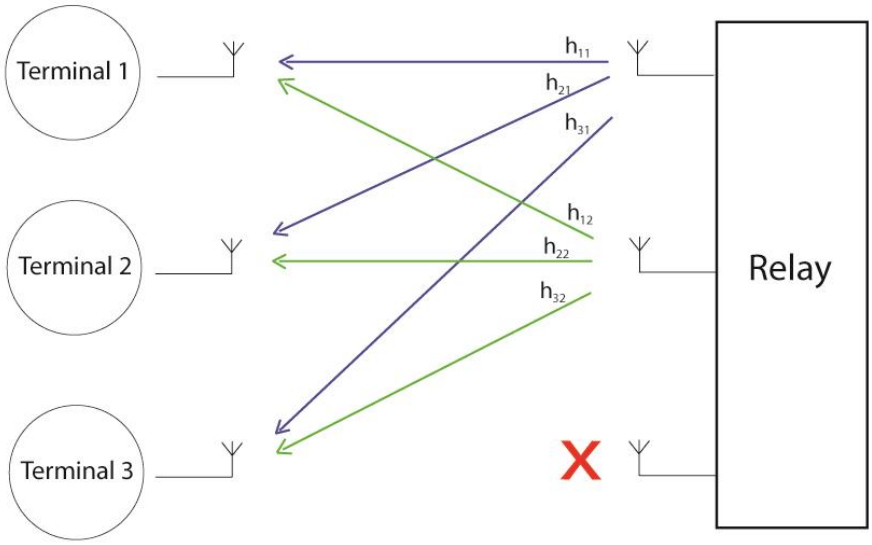

Figure 42 – First time-slot for the downlink for three time-slot strategy.

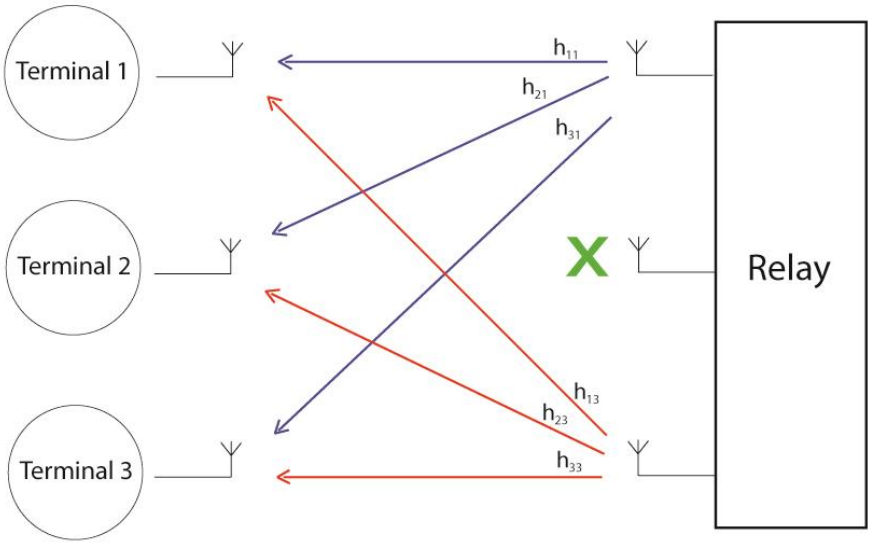

Figure 43 – Second time-slot for the downlink for three time-slot strategy.

The above-mentioned transmitting configuration and the one considered for the second time-slot of the downlink phase are assumed to be known in advance by the terminals so on the second time-slot, which corresponds to the first one in the downlink phase, the relay transmits the messages x_1 and x_2 on antenna 1 and 2, respectively. Taking into account the model described in 4.1, the received signal at each terminal on the first time-slot of the downlink phase is given by

$$y_1[2] = h_{11}x_1 + h_{12}x_2 + n\,, \tag{56}$$

$$y_2[2] = h_{21}x_1 + h_{22}x_2 + n\,, \tag{57}$$

$$y_3[2] = h_{31}x_1 + h_{32}x_2 + n\,, \tag{58}$$

where the coefficients h are the ones represented in Figure 42 and $y_{\mathrm{i}}[t]$ denote the received signal at the i^{th} terminal at the time-slot t. Notice that each terminal possesses CSI (i.e., CSIT).

At this point both terminal 1 and 2 know that they are receiving a signal that contains the message that they sent in the uplink phase so the first step of the decoding process is to cancel its own contribution from the received signal, resulting in

$$y_1'[2] = y_1[2] - h_{11}x_1\,, \tag{59}$$

$$y_2'[2] = y_2[2] - h_{22}x_2\,, \tag{60}$$

and making use of simple matched filter terminal 1 estimates $\hat{\mathrm{x}}_2$ and terminal 2 estimates $\hat{\mathrm{x}}_1$and Terminal 3 saves the received signal in order to process it later.

$$\hat{x}_2 = \mathrm{Q_C}\left[h_{12}^{-1}\, y_1'[2]\right] \tag{61}$$

$$\hat{x}_1 = \mathrm{Q_C}\left[h_{21}^{-1}\, y_2'[2]\right] \tag{62}$$

On the third time-slot, time-slot 2 of downlink Figure 43, the relay transmits the messages x_1 and x_3 on antenna 1 and 3, respectively, and the signals received by the terminals are

$$\mathrm{y}_1[3] = \mathrm{h}_{11}\mathrm{x}_1 + \mathrm{h}_{13}\mathrm{x}_3 + \mathrm{n}\,, \tag{63}$$

$$\mathrm{y}_2[3] = \mathrm{h}_{21}\mathrm{x}_1 + \mathrm{h}_{23}\mathrm{x}_3 + \mathrm{n}\,, \tag{64}$$

$$\mathrm{y}_3[3] = \mathrm{h}_{31}\mathrm{x}_1 + \mathrm{h}_{33}\mathrm{x}_3 + \mathrm{n}\,. \tag{65}$$

By now terminal 1 and 3 know that they are receiving their uplink original messages and terminal 2 can use the estimated $\hat{x}_1$. Now each terminal cancels the known contribution from the received signal, resulting in

$$y_1'[3] = y_1[3] - h_{11}x_1\,, \tag{66}$$

$$y_2'[3] = y_2[3] - h_{21}\hat{x}_1\,, \tag{67}$$

$$y_3'[3] = y_3[3] - h_{31}x_1\,. \tag{68}$$

and again using a simple matched filter both, terminal 1 and 2 estimate $\hat{x}_3$, and terminal 3 estimates $\hat{x}_1$.

$$\hat{x}_3 = Q_C\left[h_{13}^{-1}\, y'_1[3]\right] \tag{69}$$

$$\hat{x}_3 = Q_C\left[h_{23}^{-1}\, y'_2[3]\right] \tag{70}$$

$$\hat{x}_1 = Q_C\left[h_{31}^{-1}\, y'_3[3]\right] \tag{71}$$

Terminal 1 and 2 have all the messages and terminal 3 can use the message saved from the previous time slot and use $\hat{x}_1$ to apply a simple matched filter in order to estimate $\hat{x}_2$.

$$y_3'[2] = y_3[2] - h_{31}\hat{x}_1 \tag{672}$$

$$\hat{x}_2 = Q_C\left[h_{32}^{-1}\, y_3'[2]\right] \tag{73}$$

This strategy was simulated for 4 and 16-QAM modulation and the results are presented in Figure 44 and Figure 45. This performance accomplishes better results than the approach using just two time-slots described in 4.3. It is clearly less complex than the previous one since only simple computation processes and a simple matched filter are applied. However this strategy loses 50% of throughput improvement when compared with the two time-slot strategy but it still achieves a throughput improvement of 100% when compared with the traditional scheme, reducing the communication from 6 time-slots to 3 time-slots.

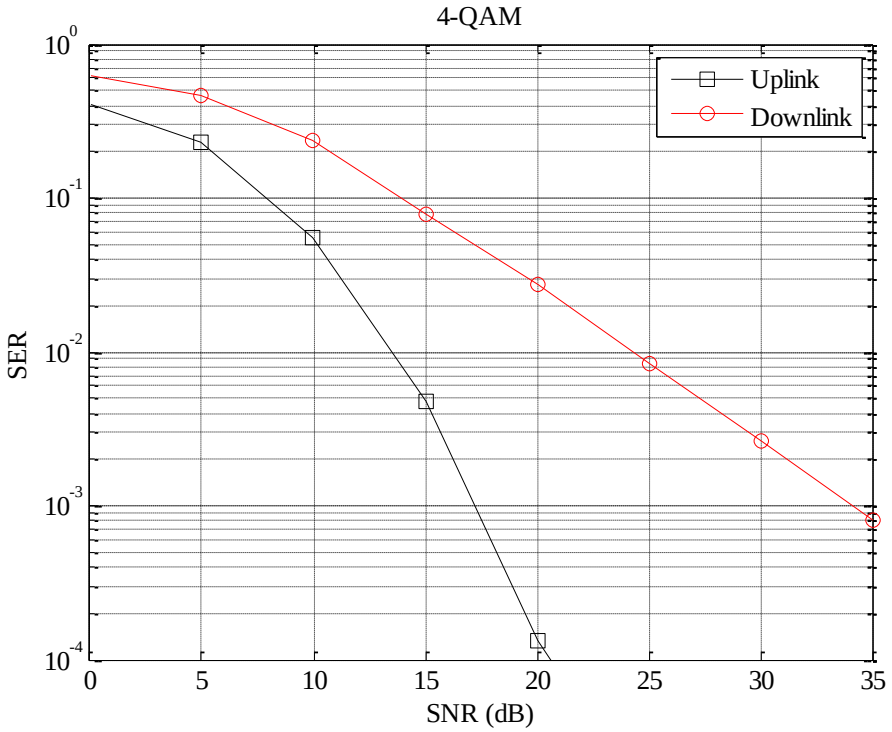

Figure 44 – 4-QAM detection in three time-slot strategy.

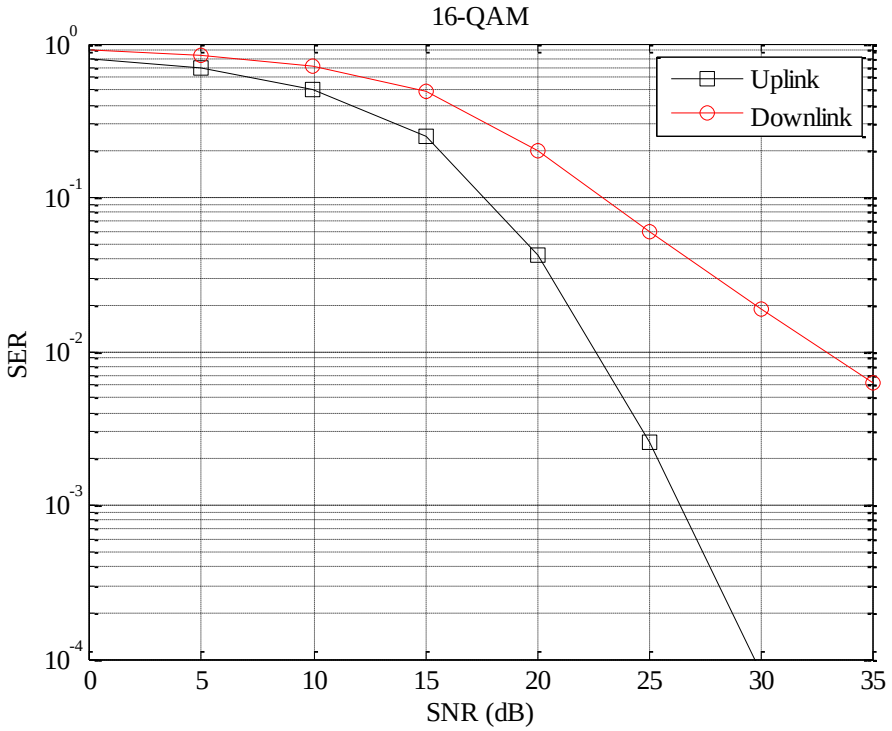

Figure 45 – 16-QAM detection in three time-slot strategy.

Section 4.5 – PLNC with MIMO relay and terminals, 2time slots strategy

The goal of this strategy is to maintain the throughput improvement of 150% over the traditional scheme, gain performance and reduce complexity in the detection process taking advantage from a new dimension. The scenario considered for the uplink stays the same specified for section 4.3.

Since one of the problems in 4.1 was the loss of one dimension on the downlink a new dimension can be gain by assuming that terminals are now equipped with two antennas instead of a single one. The uplink remains as it was considered in Figure 38, so the users transmit their messages at the same time-slot using a single antenna as previously stated.

The reception on the downlink phase makes use of the two antennas of each terminal. Since the process of the detection will be the same for the three terminals it will only be reproduced the procedure for terminal 1 Figure 46.

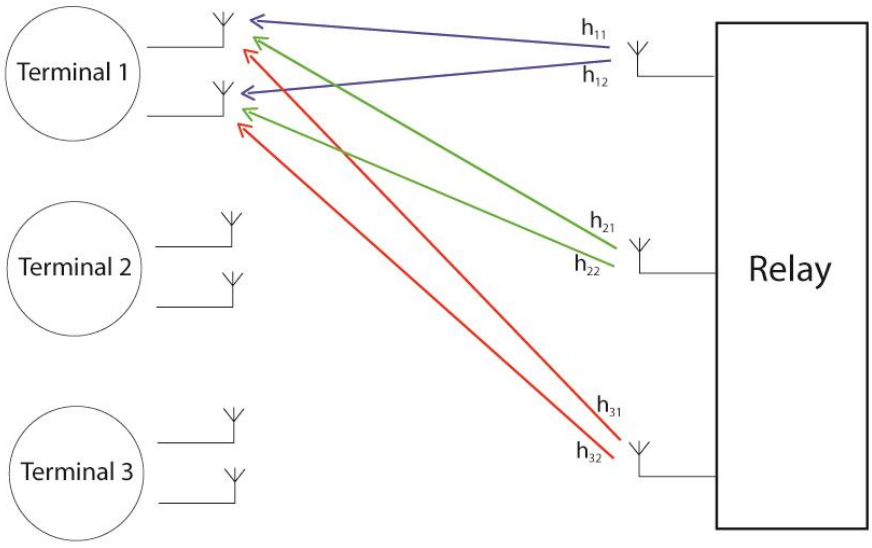

Figure 46 – Downlink for two time-slot strategy with MIMO terminals.

The MIMO relay decodes the uplink messages x_1, x_2 and x_3 then each antenna of the relay transmits one of the received messages all at the same time on the second time-slot. The sum of the received signal at the two antennas of each terminal, in this example terminal 1, is given by

$$y_{11} = h_{11}x_1 + h_{21}x_2 + h_{31}x_3 + n\,, \tag{74}$$

$$y_{12} = h_{12}x_1 + h_{22}x_2 + h_{32}x_3 + n\,. \tag{75}$$

The coefficients h are the ones represented in Figure 46 and y_{ij} denote the received signal at the i^{th} terminal and its j^{th} antenna. Since the terminal possesses CSI and knows what was its own sent uplink message the first step of the decoding process is to cancel its own contribution from the received signal, resulting in

$$y'_{11} = y_{11} - h_{11}x_1, \tag{76}$$

$$y'_{12} = y_{21} - h_{12}x_1. \tag{77}$$

After removing this self-interference cancelation the system can be rewritten as:

$$\begin{bmatrix} y'_{11} \\ y'_{12} \end{bmatrix} = \begin{bmatrix} h_{21} & h_{31} \\ h_{22} & h_{32} \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + [\mathbf{n}]\,, \tag{78}$$

$$\mathbf{y}' = \mathbf{H}'\mathbf{x}' + \mathbf{n}\,. \tag{79}$$

According with the model described in 4.1 this system is equivalent so it can be interpreted as a MIMO communications of $\underline{N_T = N_R = 2}$ antennas. Consequently all the detections strategies considered along Chapter 3 can be applied.

Summarizing in the two time-slots strategy with MIMO terminals each one of them exploits the linear codification induced by the wireless medium to reduce the detection problem to a 2×2 MIMO equivalent system and then apply a detection strategy decoding the messages from the other two terminals. This strategy is simulated for 4 and 16-QAM modulation. As expected the performance of this scheme brings down the SER curve especially for the LRA decoders. Depending on the computation capacity of the terminals it can be chosen an appropriate decoding strategy requiring more or less computation and

consequently more or less energy consumption. Within this perspective it will be presented the performances for OSIC-MMSE Figure 47 and Figure 48 and LRA OSIC MMSE Figure 49 and Figure 50 decoders once they are the ones that achieve a better performance and have a substantial difference of computation. This strategy also achieves a throughput improvement of 150% when compared with the traditional scheme, reducing the communication from 6 time-slots to 2 time-slots.

Concluding PLNC combined with a MIMO relay and MIMO users guarantees the best performance, ensuring a polynomial computational time and having a throughput improvement of 150% compared with the traditional scheme. This strategy can be generalized for scenarios with N terminals equipped with $N-1$ antennas and a MIMO relay equipped with N antennas. For instance in a scenario with four terminals the throughput improvement raises up to 200% when comparing with the traditional TDMA scheme, reducing the communication from 8 time-slots to 2 time-slots.

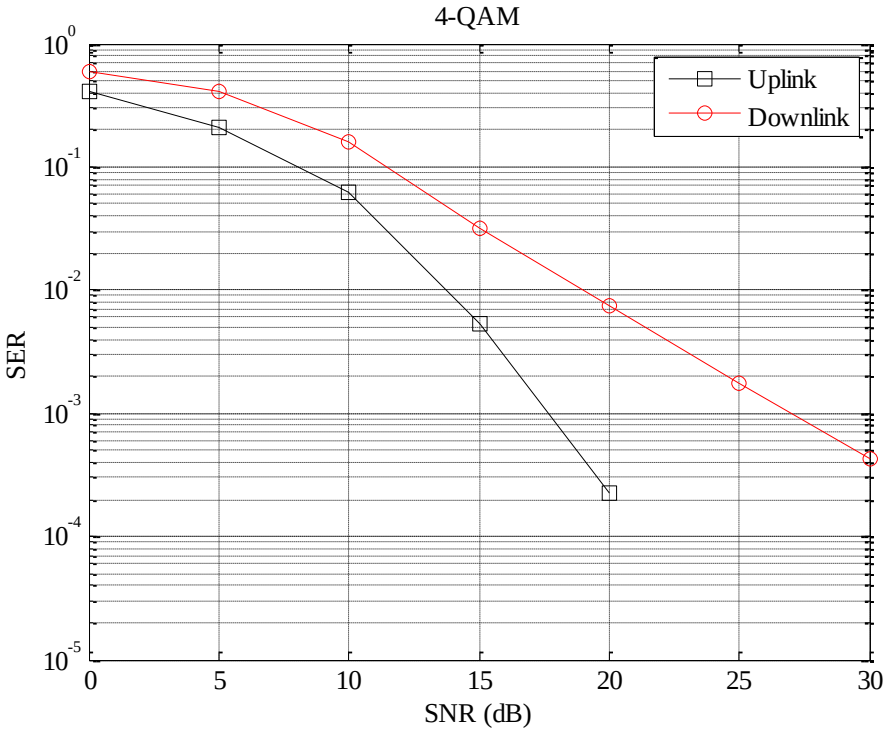

Figure 47 – 4-QAM detection in two time-slot strategy with MIMO terminals using OSIC-MMSE.

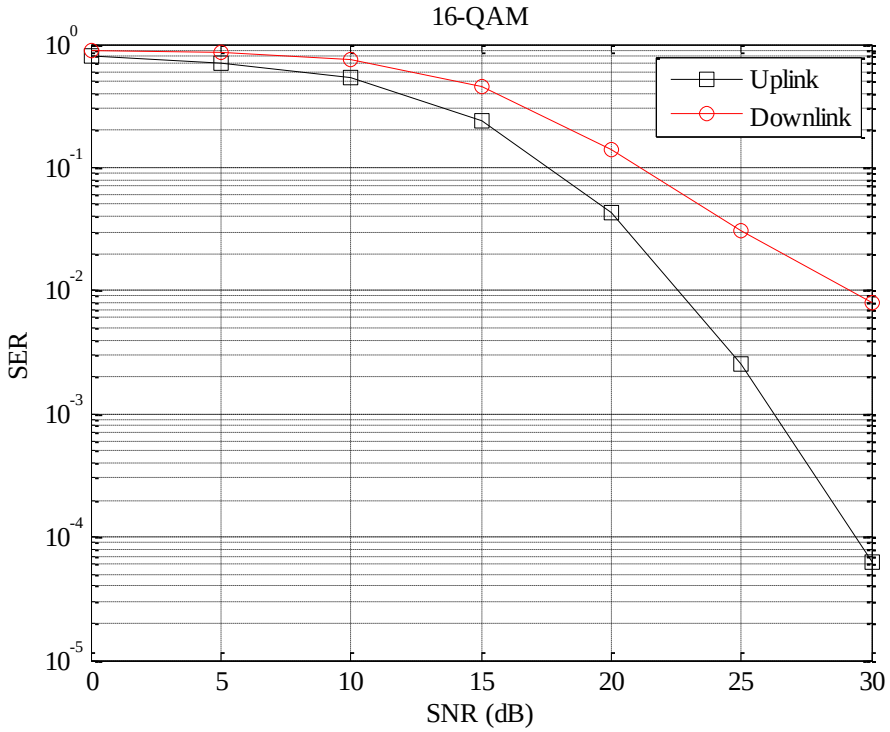

Figure 48 – 16-QAM detection in two time-slot strategy with MIMO terminals using OSIC-MMSE.

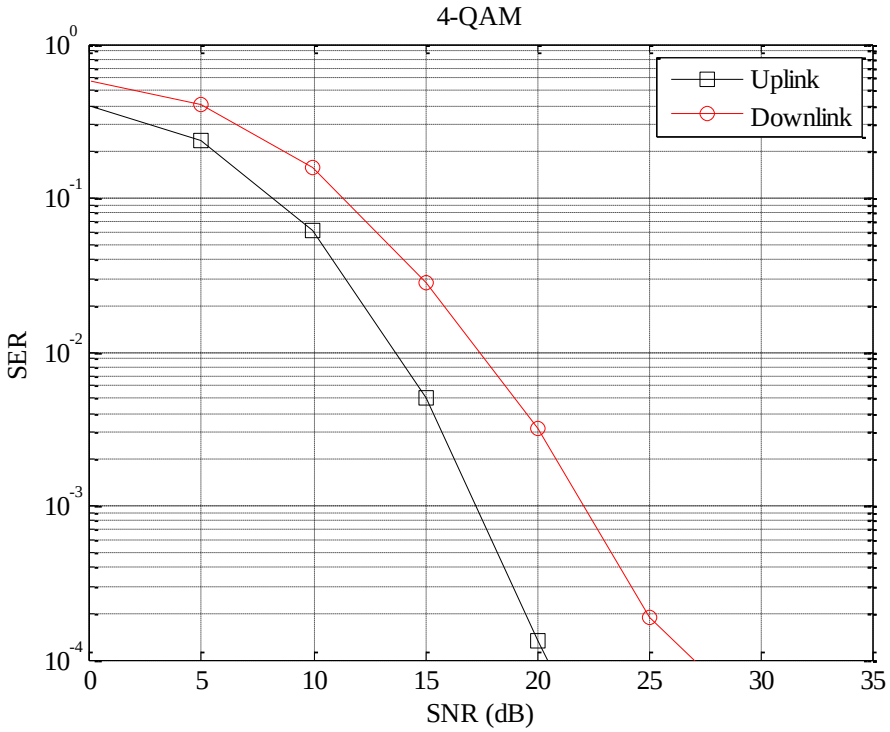

Figure 49 – 4-QAM detection in two time-slot strategy with MIMO terminals using LRA OSIC-MMSE.

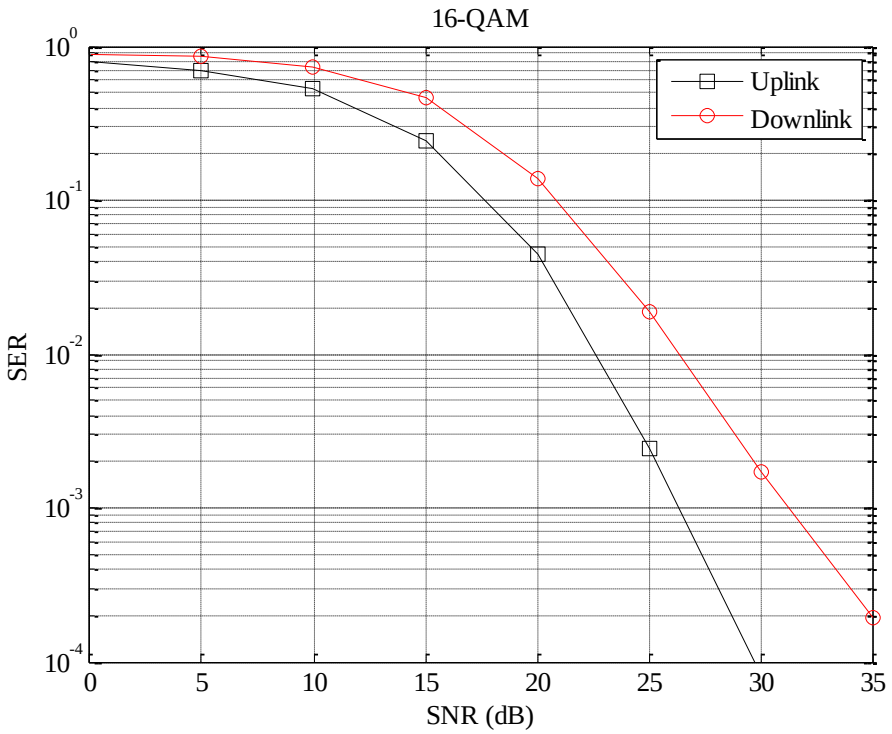

Figure 50 – 16-QAM detection in two time-slot strategy with MIMO terminals using LRA OSIC-MMSE.

Section 4.6 – PLNC with MIMO and Full Duplex antennas

The final proposal in this thesis leverages on the proposals made so far and links it with the current hot topic of full duplex (FD) radios. The next strategy has the objective of reducing all communication stages to a single time slot per message exchanges (when the messages exchanged go into infinity and the initial transient becomes negligible). by using FD terminals and loop-interference cancelation (LIC). To better explain the idea behind FD PLNC Figure 51 shows an example using TWRC system [35].

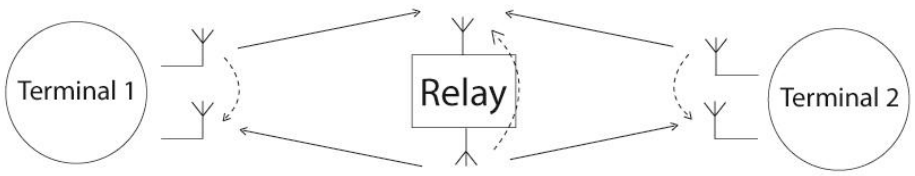

Figure 51 – Full duplex system in TWRC.

The first MAC stage is similar to the one presented in section 4.3, however, now both the relay and the terminals have extra antennas to support FD. The relay has 3 receiving antennas and 3 emitting antennas and each terminal have 1 receiving and 1 emitting antenna. This can be observed on Figure 52 for the downlink phase and Figure 53.for the uplink phase, both in the same time-slot. The main difference from the previous strategies that utilizes a single antenna on the terminal with two time slots is the loop-interference represented by the dotted lines.

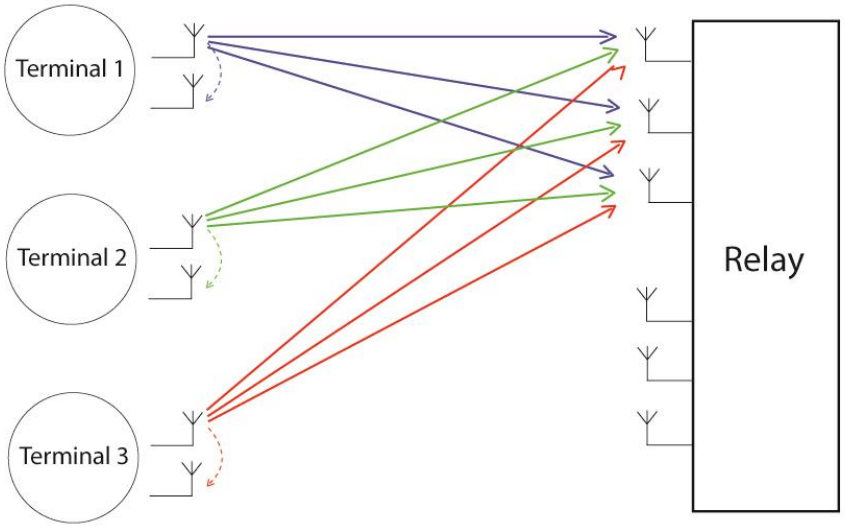

Figure 52 – Uplink phase with full duplex.

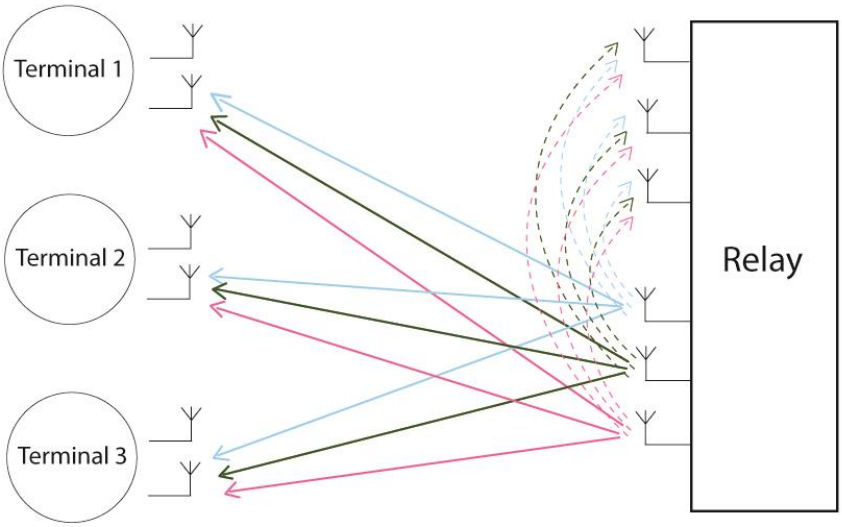

Figure 53 – Downlink phase with full duplex for the same time-slot as Fig. 52.

To better illustrate the messages changed on a single time-slot can be observed in Figure 54, where x_{ij} represents a message sent from terminal *i* during time-slot *j*.

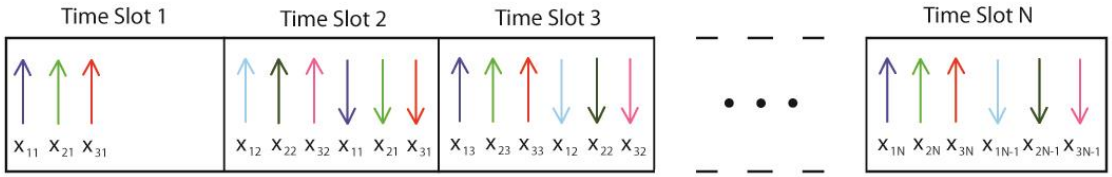

Figure 54 – Messages exchange between the terminals and the relay.

Since on the first time-slot the relay doesn't have anything to send to the terminal it remains silent, after that initial time-slot all the transmitter antennas are sending messages.

The following examples will be on all transmissions after the first. The signals that the relay receives is

$$\mathbf{y}_2 = \mathbf{H}_2\mathbf{x}_2 + \mathbf{n}_2 + \mathbf{H}_{\mathrm{Li}}\hat{\mathbf{x}}_1 \tag{80}$$

where $\mathbf{y}_2$ is the received signal in the second time slot, $\mathbf{H}_{\mathrm{Li}}$ is the channel of the LI, $\hat{\mathbf{x}}_1$ is the previous detected symbol at the relay and $\mathbf{H}_2\mathbf{x}_2$ is the sent messages by the terminal passing through the channel.

In each terminal it is very similar but instead of the LI being with the previous symbol is with the next one

$$\mathbf{y}_2 = \mathbf{H}_2\hat{\mathbf{x}}_2 + \mathbf{n}_2 + \mathbf{H}_{\mathrm{si}}\mathbf{x}_3 \tag{81}$$

As we can observe the LI contribution is very big and can lead to many errors, so some type of isolation (physical or electrical) has to be added between the emitting and receiving antennas so this new term k will be added to our simulation to reduce the SI. Throughout the rest of this chapter the factor *K* will take the values of 10^{-1}, $5\text{x}10^{-2}$, 10^{-2} and 10^{-3} (i.e., –10 dB, –13 dB, –20 dB, and –30 dB respectively). Therefore the previous expressions will become

$$\mathbf{y}_2 = \mathbf{H}_2\mathbf{x}_2 + \mathbf{n}_2 + K\mathbf{H}_{\mathrm{si}}\hat{\mathbf{x}}_1 \tag{82}$$

$$\mathbf{y}_2 = \mathbf{H}_2\hat{\mathbf{x}}_2 + \mathbf{n}_2 + K\mathbf{H}_{\mathrm{si}}\mathbf{x}_3 \tag{83}$$

Figure 55 and Figure 56 correspond to the simulation results using 4 and 16-QAM modulations for all the values of *K*.

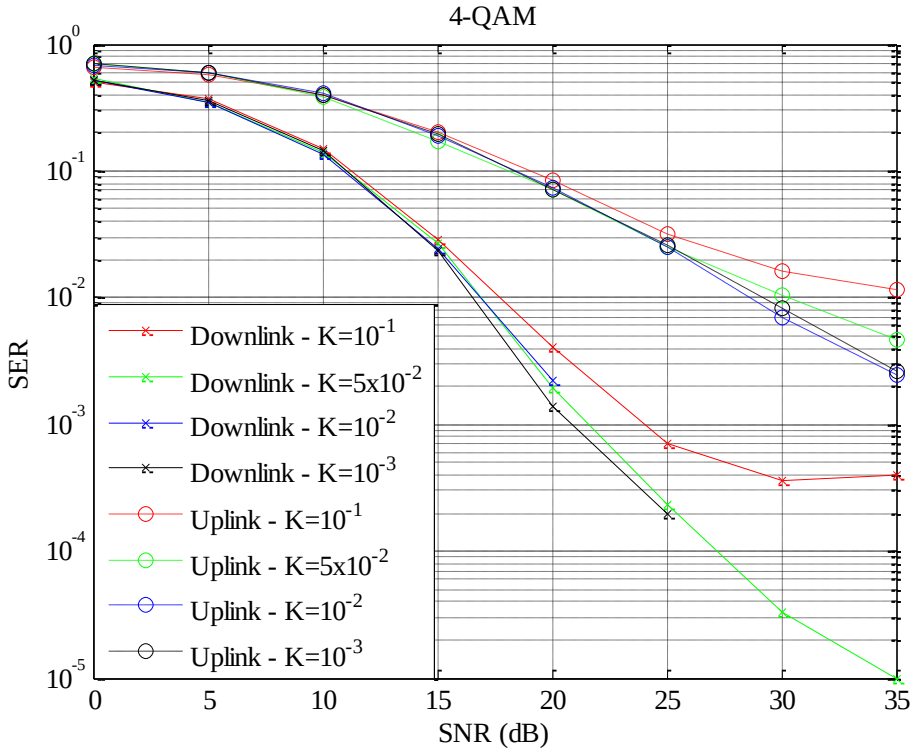

Figure 55 – 4-QAM detection in two time-slot strategy using full duplex.

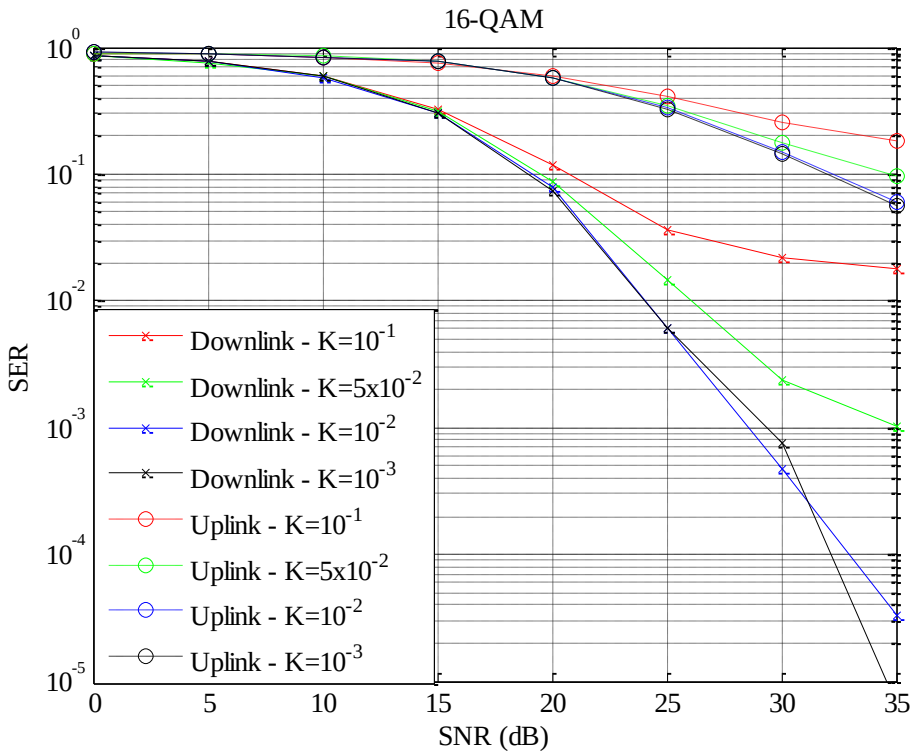

Figure 56 – 16-QAM detection in two time-slot strategy using full duplex.

The second simulation is like the one presented on section 4.5 but the relay has 3 receiving antennas and 3 emitting antennas and each terminal have 2 receiving and 1 emitting antenna. This can be observed on Figure 57 for the downlink phase and Figure 58.

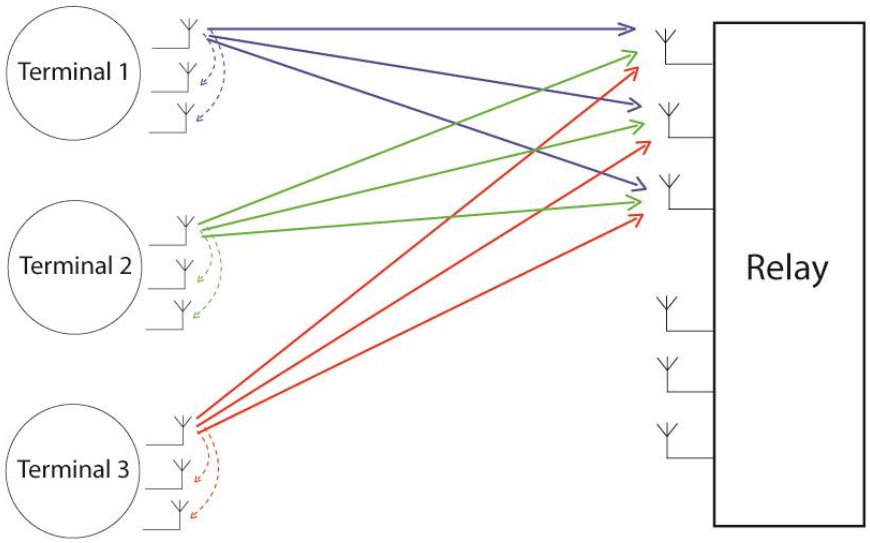

Figure 57 – Uplink phase with full duplex and MIMO terminals.

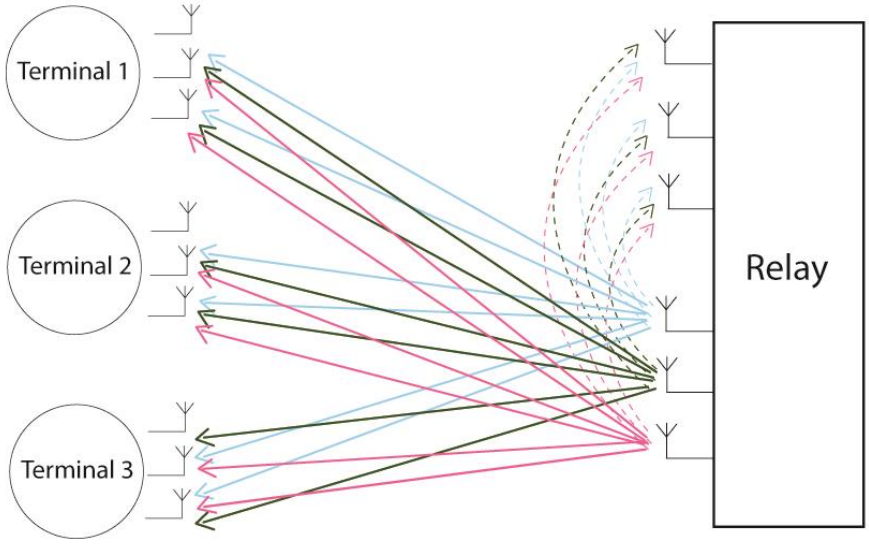

Figure 58 – Downlink phase with full duplex and MIMO terminals for the same time-slot as Fig. 55.

Figure 59 and Figure 60 correspond to the 4 and 16-QAM modulations simulation using all the values of k.

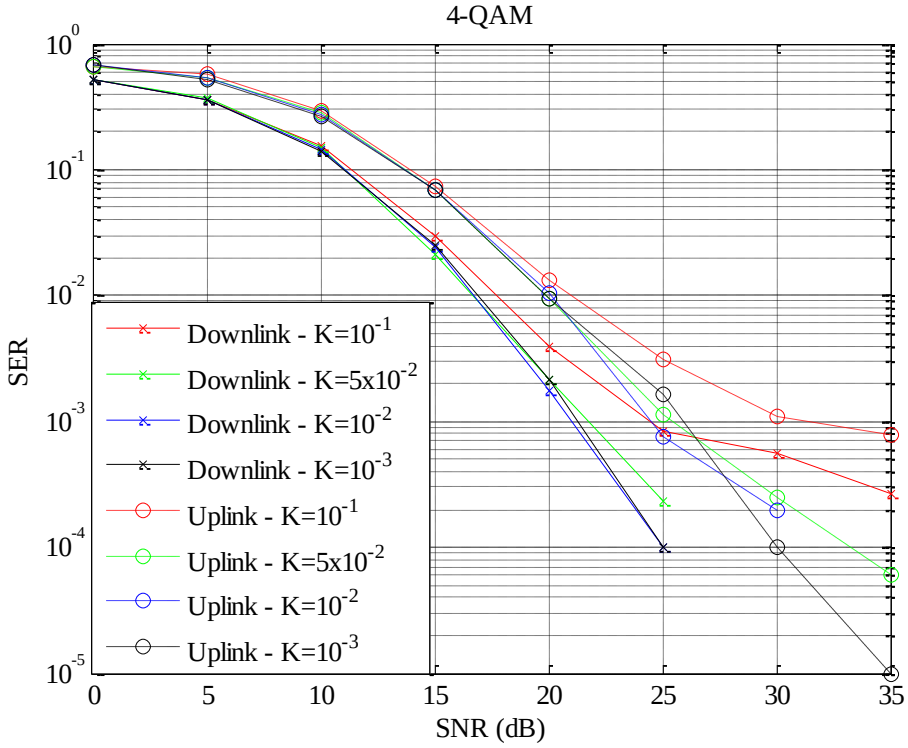

Figure 59 – 4-QAM detection in two time-slot strategy using full duplex and MIMO terminals.

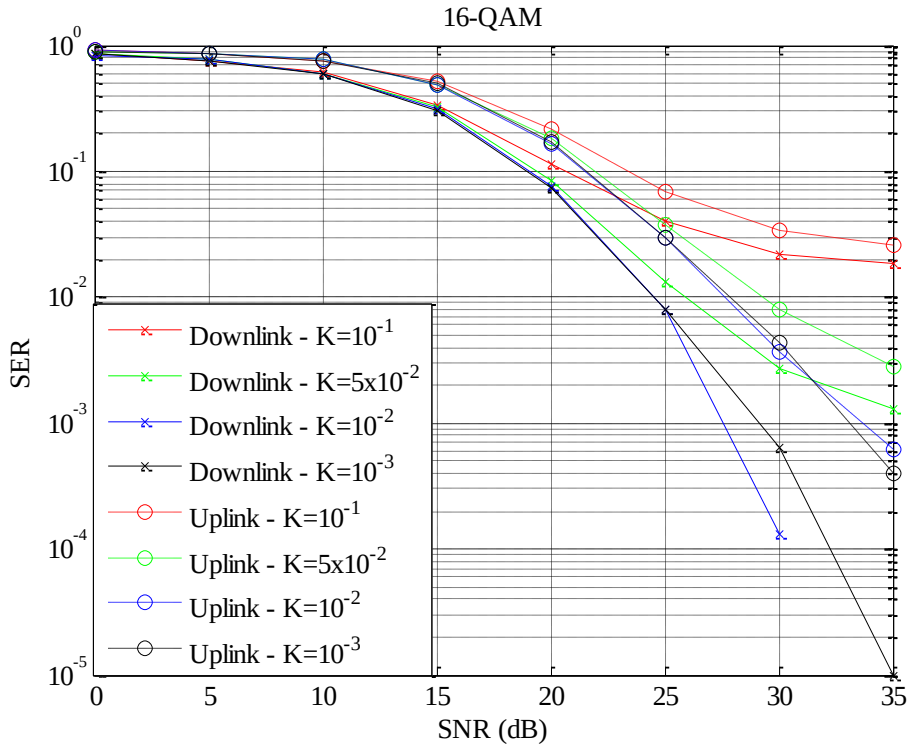

Figure 60 – 16-QAM detection in two time-slot strategy using full duplex and MIMO terminals.

One can observe that the higher the *K* value is the worst results we get in all simulations, this is expected since *K* is the isolation. Both sistems proposed have similar results to the ones before that didn't use full-duplex for lower values of *K* which proves that with the proper isolation full-duplex is a good strategy to implement.

Chapter 5

Conclusions

The last Chapter points to the conclusions of this dissertation as well as possible future direction for the work developed

Section 5.1 – Main conclusions

This dissertation presented several strategies for the multi-way channel combining principles of physical-layer network coding with MIMO spatial multiplexing and full duplex systems, which lead to a considerable reduction in the number of time-slots necessary to exchange information between all terminals through a relay.

The manipulation of geometric lattices played a central role in this dissertation given that the signals we dealt with in this work all come from discrete alphabets and when they are transmitted, one always gets linear combinations of those symbols, which, by construction, always define a (bound/finite) lattice.

A number of detectors where explored in Chapter 3 in order to explain how they find solutions to the closest vector problem in MIMO systems. The results achieved show that, although LRA receivers capture the same diversity order of ML detection (i.e., $d=N$), the LRA OSIC-MMSE receiver is the one that achieves the best performance since it is the one that has the smaller power penalty. Moreover, instead of an exponential time complexity as in the ML case, it rather has a polynomial time complexity, mostly because the LLL algorithm has polynomial complexity. However, both these detectors r imply the need for large computational complexity, and therefore they turn out not very efficient for using in the terminals. For that reason, sub-optimal solution have to be adopted. In general ZF, MMSE and OSIC have the same diversity order, $d=1$, but OSIC presents the highest gain. With a higher number of antennas or higher order M-QAM constellations, ZF and MMSE increasingly degrade their performance in comparison to ML.

The concepts of PLNC that where introduced in Chapter 1 were put in practice in Chapter 4 where MIMO spatial multiplexing was combined with PLNC in scenarios others than the traditional TWRC. The strategies presented use distributed MIMO in the uplink phase and different approaches were developed and assessed for the downlink phase.

One strategy makes use of two time-slot with a single antenna and for that reason it leads to the lowest performance of them all. This can be explained by the fact that ML detection is implemented using the projection of signals in just one dimension rather than the two dimension needed to decide for the symbols (and besides that, its complexity is exponential). Nevertheless, it has a 150% improvement on the throughput when compared with the traditional TDMA.

The three time-slot strategy has a better performance than the previous one because it captures the information in the full dimensions of the incoming signals, however, it only achieves a 100% throughput increase when compared to TDMA.

The two time-slots strategy making use of multiple antennas in the terminals is the strategy achieving the best performance while maintaining a 150% throughput improvement in respect to TDMA and at the same time doing that in polynomial time. Moreover, any of the sub-optimal detectors described on Chapter 3 can be used in this strategy. Among all the detectors, the one proving to have the best performance was LRA OSIC-MMSE, as known in the literature.

The last strategy proposed in this thesis adds full duplex capability two the previous strategies: the two time-slots with single antennas and the two time-slots with MIMO antennas. The main advantage in this strategy it is not the computation need, but rather a 200% increase throughput when compared to TDMA. In this scenario the terminals are simultaneously receiving and transmitting messages in the same frequency band, which are able to accomplish all the exchanges of message in one time-slot only on average (as the number of exchanges tends to infinity).

It is worth noting that both strategies that use MIMO antennas can be generalized to scenarios with N terminals, each one equipped with $N-1$ antennas and the relay having N antennas in the case of half-duplex, and $2N$ antennas in the case of full duplex.

Section 5.2 – Future work

One research line that should be explored is to expand the two time-slot strategies (both single antennas and MIMO cases) to *N* terminals.

In [36] the authors suggest multipar decode-and-forward relay channel, using full-duplex in a relay station equipped with a massive arrays of antennas in order to reduce loop interference. This novel strategy could be implemented and be improved in the near future.

Lattices were one important tool in this dissertation because of their role in understanding how MIMO detectors work**.** As presented in Chapter 1, another branch of PLNC is compute-and-forward, where lattices are the central tool for code construction given their group property and the isomorphism that exists between lattice structures and group codes (whose words are associated with lattice points), given that the sum of any codewords will always be another valid codeword. This branch of algebraic PLNC is very promising for further efficient use of wireless networks resources.

Bibliography

[1] S. Y. R. Li, R. W. Yeung, Ning Cai, "Linear network coding", *IEEE Trans. on Information Theory*, vol. 49, no. 2, pp. 371 – 381, February 2003

[2] S. Zhang, S. C. Liew, P. P. Lam, "Hot topic: Physical-layer Network Coding", in *Proc. of ACM MobiCom'06*, pp. 358-365, Sept 2006.

[3] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network", in *Proc. of IEEE Int. Conf. on Communications*., Istanbul, Turkey, pp. 3885-3890, June 2006.

[4]B. Nazer and M.Gastpar, "Computing over multiple-access channels with connections to wireless network coding", in *Proc. of IEEE Int. Symp. on Inf. Theory*, Seattle, Washington, USA, pp. 1354-1358, July 2006.

[5] B. Nazer and M. Gastpar, "Compute-and-forward: A novel strategy for cooperative networks", in *Proc. of the 42nd Asilomar Conference on Signal, Systems and Computers*, October 2008.

[6] B. Nazer and M. Gastpar, "Reliable Physical Layer Network Coding", *Proceedings of the IEEE*, vol. 99, no. 3, March 2011.

[7] Chen Feng, Danilo Silva and Frank R. Kschischang, "An algebraic approach to physical layer network coding", in *Proc. of ISIT'10 – The Inter. Symp. on Information Theory*, Austin, Texas, USA, pp. 1017-1021, June 2010.

[8] P. Q. Nguyen and B. Vallée, "The LLL algorithm (information Security and Cryptography", Springer Berlin Heidelberg; 1 edition ,December 2009

[9] Daniele Micciancio, *Lattices Algorithms and Applications*, University of California, San Diego, Winter 2010

[10] Oded Revev, "Lattices in Computer Science", *Lecture notes*, Tel Aviv University, Fall 2004.

[11] A. K. Lenstra, H. W. Lenstra and L. Lovász, "Factoring polynomials with rational coefficients", *Math. Annals*. vol. 261, no. 4, pp. 515-534, 1982

[12] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equations", *IEEE Transactions on Communications*, vol. 56, no. 2, pp. 309-318, February 2008

[13] D. Tse and P. Viswanath, *Fundamentals of wireless Communications*, Cambridge University Press, New York, USA, 2005.

[14] A. J. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmissions/directional reception", *Technical Report U.S. Patent*, no. 5,345,599 , 1994

[15] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Syst. Tech. J*., pages 41–59, Autumn 1996.

[16] G. J. Foschini, G. D. Golden, R. A. Valenzuela and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communications employing multi-element arrays", *IEEE J. Select. Areas Commun*. 17:1847-1852, November 1999.

[17] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*., 6: 311–335, March 1998.

[18] P. Viswanath, D. N. C. Tse, and V. Anantharam, "Asymptotically optimal waterfilling in vector multiple-access channels," *IEEE Trans. On Inf. Theory*, vol. 47, no. 1, pp. 241–267, January 2001.

[19] A. Sibille, C. Oestges and A. Zanella, "MIMO from theory to implementation", *Academic Press*, Elsevier, 2011.

[20] R. C. de Lamare, "Massive MIMO systems. Signal processing challenges and research trends", *IEEE Trans. on Information Theory*, Submitted on 28 October 2013.

[21] C. Windpassinger, "Detection and precoding for multiple input output channels", *PhD dissertation*, University of Erlangen-Nürnberg, Erlangen, Germany, 2004.

[22] P. Q. Phong and B. Vallée, *The LLL algorithm: survey and applications*, Heidelberg: Springer, 2010.

[23] L. Babai, "On Lovasz' lattice reduction and the nearest lattice point problem", *Combinatorica*, vol. 6, no.1, pp. 1-13, 1986.

[24] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems", in *Proc. of Globecom – IEEE Global communications Conf*, Taipei, China, November 17-21, 2002.

[25] C. Ling, "Towards Characterizing the Performance of Approximate Lattice Decoding in MIMO Communications",in *Proc. of the 4th Inter. Symp. on Turbo Codes & Related Topics*, April 2006.

[26] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, volume I, John Wiley & Sons, Inc., New York, NY, USA, 1968.

[27] F. A. Monteiro, "Lattices is MIMO spatial multiplexing: Detection and geometry", *PhD dissertation*, Department of Engineering, University of Cambridge, May 2012.

[28] G. D. Golden, C. J. Foschini, R. A. Valenzuela and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-Blast space-time communication architecture", *IET Electronics Letters*, vol. 35, no. 1, January 1999.

[29] A. Gorokhov, D. Gore and A. Paulraj, "Performance bounds for antenna selection in MIMO system Communications", in *Proc. Of IEEE ICC Int. Conf. on Communications*, vol. 5, 3021 – 3025, 2003.

[30] B. Nazer and M. Gastpar, "Compute-and-forward: A novel strategy for cooperative networks", in *Proc. of 42nd Asilomar Conference on Signals, Systems and Computers*, October 2008.

[31] A. Goldsmith, "Wireless Communications", *Cambridge Univ. Press*, 2005.

[32] M. Morelli, "Timming and frequency synchronization for the uplink of an OFDMA system", *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296 – 306, February 2004.

[33] N. Xu, "Physical-layer network coding for MIMO systems", *PhD dissertation*, Computer Science and Engineering Department, University of North Texas, May 2011.

[34] N. Lee, J. Lin and J. Chun, "Degrees of Freedom of the MIMO Y channel: Signal Space Alignment for Network Coding" ", *IEEE Trans. on Information Theory*, vol. 56, no. 7, pp. 3332–3342, July 2010

[35] S. Tedik and G. Kurt, "Practical full-duplex physical layer network coding", *accepted in VTC*, 2014

[36] H. Q. Ngo, H. A. Suraweera, M. Matthaiou and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE Journal on Selected Areas in Communications, special issue on Full-duplex Wireless Communications and Networks*, vol. 32, pp. 1721-1737, Sept. 2014