Complexity improved Sphere Decoder for Highly Correlated and LOS channels

Juan Francisco Castillo León, Rafael Aguilar González, Ulises Pineda Rico, and Enrique Stevens-Navarro

Facultad de Ciencias
Universidad Autónoma de San Luis Potosí
Salvador Nava S/N, Zona Universitaria
S.L.P., Mexico
e-mail:jncastillo45@gmail.com, rafa_ag108@hotmail.com,
u_pineda@galia.fc.uaslp.mx, estevens@galia.fc.uaslp.mx

Abstract. In recent years it has been shown that iterative decoding techniques improve the performance (bit error rate) of various digital communication systems. Techniques for Multiple-input multiple-output (MIMO) are a key technology to promote high-speed wireless communication under the need of a low complexity iterative scheme for detection. Spherical decoding (SD) has been suggested as an efficient algorithm to solve such detection problem. SD is known as an algorithm of polynomial complexity without clearly specifying the assumptions made about its structure. Recently, SD has become a powerful tool to achieve a performance close to maximum likelihood (ML) algorithm (considered ideal) but involving lower complexity. In this paper we analyze the spherical decoder performance compared to other decoders on different channels using Montecarlo simulations.

Keywords: Sphere Decoding, Wireless Communications, Multi-Antenna Systems, Maximum Likelihood, Zero Forcing, Complexity.

1 Introduction

Wireless communications have captured the attention and imagination of the world and have become the segment's largest and fastest growing subject in the area of telecommunications. The main reasons are the desire for mobility and access to the network without a physical connection (wired). Various technologies and systems have been proposed to provide wireless communication services [1].

The first generation mobile systems (1G) is characterized by analog transmission of voice and it was introduced in the early 80's. Systems of second generation (2G) are distinguished by the digital transmission of voice and data at low rates. The third generation (3G), arises from the need to increase data transmission capacity in order to offer services like Internet access, video conferencing, television, and downloading files [2].

The success of wireless communications has been primarily associated with a steady increase in system capacity and quality of service. The bandwidth is limited and expen-

Advances in Computer Science and Electronic Systems
Research in Computing Science 52, 2011, pp. 308-316
Received 23/01/11
Accepted 20/02/11
Final version 07/03/11

sive and in order to continue this trend should be used new technologies to provide greater spectral efficiency and reliability.

Traditionally the antenna systems are formed by one transmitter and one receiver, e.g. Single-Input Single-Output (SISO) systems. In some cases this setting is changed by increasing the number of antennas to provide diversity to the system, transforming it in a Multiple-Input Multiple-Output (MIMO) system. The objective of MIMO systems is to increase the capacity given the rich scattering propagation environment offered to the signal.

The work of Foschini [3] and Telatar [4] show that by increasing the number of antennas on both sides of the channel, substantially increases the number of bits that can be transmitted (capacity), something unthinkable in SISO systems. This increased capacity is associated with a wealth of dispersion in the environment, which allows the transmission of information by independent paths.

Due to its advantages over traditional systems, the MIMO communication systems have emerged as a key technology. MIMO techniques have been proposed as extensions of existing wireless communication standards such as IEEE 802.11, HSDPA and are part of emerging standards such as IEEE 802.16.

There are generally three categories of MIMO techniques. The first aims to improve power efficiency and maximization of spatial diversity. For example, delay diversity, STBC (Space Time Block Codes), STTC (Space Time Trellis Codes). The second approach uses layers to increase the capacity, e.g. V-BLAST (Vertical-Bell Laboratories Layered Space-Time) where signals are transmitted over multiple antennas to increase transmission speed. The third type exploits the channel knowledge at the transmitter. This channel information is used for pre and post filtering in the transmitter and receiver, which can achieve a gain in capacity.

Pre-coded data to be transmitted cannot completely prevent the effects of the channel due power constraints. Also, there are few problems with the calculation of the inverse of the channel, especially when the matrix-channel is near singular or singular. Therefore, it is necessary a stage for detection at the receiver in order to ensure successful information recovery. The detection methods can be optimum (that are often complex) or suboptimal (Heuristic) which have a low computational complexity.

The Maximum likelihood (ML) detector, in general terms, it requires joint detection of an entire block of symbols [5]. Although optimal, the extreme complexity of the decoder is opposed to practical use in multiple antenna systems. Especially, when using modulations of several bits per symbol and / or many transmit antennas are involved. For multiuser detection (MUD), the block of symbols increases and so does the number of operations needed to detect, making it virtually impossible for practical use [6]

Zero Forcing Detector (ZF) uses the reverse of the channel to remove the effects of it, but despite its low complexity is not useful for practical applications since the calculation of the inverse of the channel becomes complex by increasing its size. Additionaly, the channel matrix may be not invertible and its performance is far below the ML detector.

Consequently, there has been a growing interest in the field of decoding for ML detection in digital communications. The sphere decoding (SD), offers to decrease the computational complexity because it only explores the possible outcomes in a radio "*r*" thus reducing the number of operations performed to obtain a result [7].

In this paper, we analyze the performance and complexity advantages of a proposed spherical decoder comparing it with others using Monte Carlo simulations.

2 System Model

The model consists of a MIMO system with M_T transmitters and M_R receiving antennas, the received signal vector of dimension M_R is given by:

$$y = Hx + n \tag{1}$$

where H denotes the channel matrix $M_R \times M_T$, $x = [x_1 \quad x_2 \quad \dots \quad x_{M_T}]^T$ is the signal transmitted vector of M_T elements and n is a complex Gaussian noise vector that is added with dimensions M_R. Inputs x are chosen independently from a constellation O to the bits per symbol Q, e.g. $|O| = 2^Q$. The set of all possible symbols to transmit is denoted by O^{M_T}. We assume for the simulation that the number of receivers equals the number of transmitters $M_T = M_R$ and also H is modeled as a Rayleigh fading channel, Rician or correlated [8].

3 Overview of Methods

3.1 Maximum Likelihood

Maxium likelihood (ML) is based on the method of least squares and the objective is to find the minimum Euclidean distance of each element from the vector received while at the same time analyzing all existing solutions, see Figure 1.

$$\widehat{x} = \arg \min_{x \in R^{M_R}} \|y - Hx\|^2 \tag{2}$$

Basically, ML consists in solving (2) from a set of possible symbols which depend entirely in the shape of the modulation scheme used. From (2) we have that $y = [y_1\, y_2\, \cdots, y_{M_R}]$ is the received vector and $H_{M_T \times M_R}$ corresponds to the channel, $x = [x_1\, x_2\, \cdots\, x_{M_T}]$ is the potential vector data that has been sent to, $\widehat{x} = [\widehat{x}_1\, \widehat{x}_2\, \cdots\, \widehat{x}_{M_R}]$ is the vector of received data according to the algorithm, which theoretically was sent. With this method, the complexity grows exponentially as it makes 2^{M_T} iterations before delivering a result.

3.2 Zero Forcing

For a channel with a response H, ZF decoder inverts the channel response calculating its inverse.

$$inv(H) = H^{-1} \tag{3}$$

$$H * H^{-1} = I \tag{4}$$

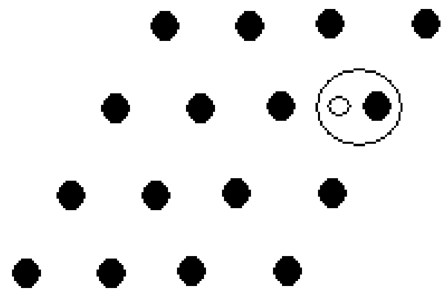

Fig. 1. Interpretation of ML.

Ideally, when performing this process, the channel effects are nullified as seen in (4).

$$y = Hx + n \tag{5}$$

$$\widehat{x} = inv(H) * y \tag{6}$$

$$\widehat{x} = x + n' \tag{7}$$

however, as shown in (7), the noise vector has been amplified by H^{-1}.

3.3 Sphere-Decoder

The basic premise of sphere decoding is quite simple: it comes to finding the minimum Euclidean distance within a sphere centered at y and radius r reducing the search space and therefore the required calculations, as shown in Figure 2.

It is clear that the closest point within the radius of the sphere is also the closest point within the full mesh.

The point Hx is a sphere of radius "r", if and only if:

$$r^2 \geq \|y - Hx\|^2 \tag{8}$$

To divide the above problem into subproblems, it is useful to consider the QR factorization of the matrix H.

$$H = QR \tag{9}$$

$$H = \begin{bmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ 0 & R_{22} & R_{23} \\ 0 & 0 & R_{33} \end{bmatrix} \tag{10}$$

So then the condition can be described as:

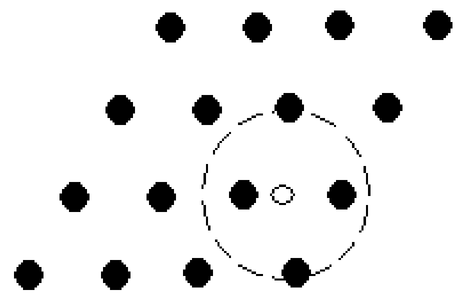

Fig. 2. Interpretation of SD.

$$r^2 \geq \|Q_1^* y - Rx\|^2 + \|Q_2^* y\|^2 \tag{11}$$

or:

$$r^2 - \|Q_2^* y\|^2 \geq \|Q_1^* y - Rx\|^2 \tag{12}$$

defining $y = Q_1^* y$ and $r'^2 = r^2 - \parallel Q_2^* x \parallel^2$ allows us to rewrite this as

$$r'^2 \geq \sum_{i=1}^{m} (y_i - \sum_{j=i}^{m} r_{i,j} x_j)^2 \tag{13}$$

where the first term depends only on x_m, the second term on x_m, x_{m-1}, and so on. Therefore a necesasary condition for Hx to lie inside the sphere is that $r'^2 \geq (y_m - R_{m,m} x_m)^2$. This condition is equivalent to x_m belonging to the interval

$$\left\lceil \frac{-r' + y_m}{R_{m,m}} \right\rceil \leq x_m \leq \left\lfloor \frac{r' + y_m}{R_{m,m}} \right\rfloor \tag{14}$$

where $\lceil \bullet \rceil$ denotes rounding to the nearest larger element in the set of numbers that spans the lattice. Similarly, $\lfloor \bullet \rfloor$ denotes rounding to the nearest smaller element in the set of numbers that spans the lattice.

Of course, (14) is by no means sufficient. For every x_m satisfying (14), defining $r'^2_{m-1} = r'^2 - (y_m - R_{m,m} x_m)^2$ and $y_{m-1|m} = y_{m-1} - R_{m-1,m} x_m$ a stronger necessary condition can be found by looking at the first two terms in (13), which leads to belonging to the interval

$$\left\lceil \frac{-r'_{m-1} + y_{m-1|m}}{R_{m-1,m-1}} \right\rceil \leq x_{m-1} \leq \left\lfloor \frac{r'_{m-1} + y_{m-1|m}}{R_{m-1,m-1}} \right\rfloor \tag{15}$$

Algorithm

Input: Q, R, $y = Q_1 x$, r.

1. (initialize) $k = m$, $r_m'^2 = r^2 - \|Q_2^* x\|^2$, $y_{m|m+1} = y_m$
2. (limits) $UB(k) = \lfloor (r_k' + y_{k|k+1})/R_{k,k} \rfloor$, $x_k = \lceil (-r_k' + y_{k|k+1})/R_{k,k} \rceil - 1$
3. (increase x_k),$x_k = x_k + 1$. If $x_k \leq UB(k)$, go to 5; else , goto 4.
4. (increase k) $k = k+1$; if $k = m+1$, terminate the algorithm; else, goto 3.
5. (Decrease k) If $k = 1$, ir a 6; else $k = k-1$, $y_{k|k+1} = y_k - \sum_{j=k+1}^{m} R_{k,j} x_j$, $d_k'^2 = d_{k+1}'^2 - (y_{k+1,k+2} - R_{k+1,k+1} x_{k+1})^2$, and goto 2.
6. Solution found. save x and it's distance from y, $d_m'^2 - r_1'^2 + (y_1 - R_{1,1} x_1)^2$ and goto 3.

Where Q and R come from the decomposition QR, y is the received data, $y = Q_1 x$, r is the radius of the sphere, m is the dimension received data vector and $\widehat{x}$ is the estimated figure [7].

4 Results

In figures 3 and 4 is observed as a reference the ZF detector performance over a Rayleigh fading channel, although the detector has a low computational complexity, its performance is inferior to the ML detector and the SD.

With a SNR of 10 dB and a Rician fading channel (which corrupts the data in a more aggressive way than the Rayleigh channel) with $k = 0.1$, we obtain a gain of 2.5 dB compared to ZF detector. However, this gain causes an increase in computational complexity, see Figure 3

Using the SD under similar conditions to those mentioned above, but this time on a correlated fading channel with $alpha = 0.5$, we obtain a performance similar to ZF. When $alpha < 0.5$ (highly correlated channel), we obtain a certain gain causing again a complexity increment, see Figure 4

The ideal solution to the problem is given by the ML scheme, but due to the exhaustive search performed along all the constellation of possible outcomes, it becomes prohibitive to practical use. In other words, complexity rises rapidly with increasing the number of antennas or change in the encoding (increase of bits per symbol). The SD reduces the complexity and try to get similar performance so it can be implemented and its complexity in the worst case is polynomial [5] making it more practical than ML.

The Figure 5 compares the computational complexity of ML detector and SD in terms of FLOPS (Floating Point Operations Per Second), we can see that increasing the number of antennas ($Tx = Rx$) and therefore the possible outcomes, the ML scheme increases exponentially the number of operations needed for the detection stage. On

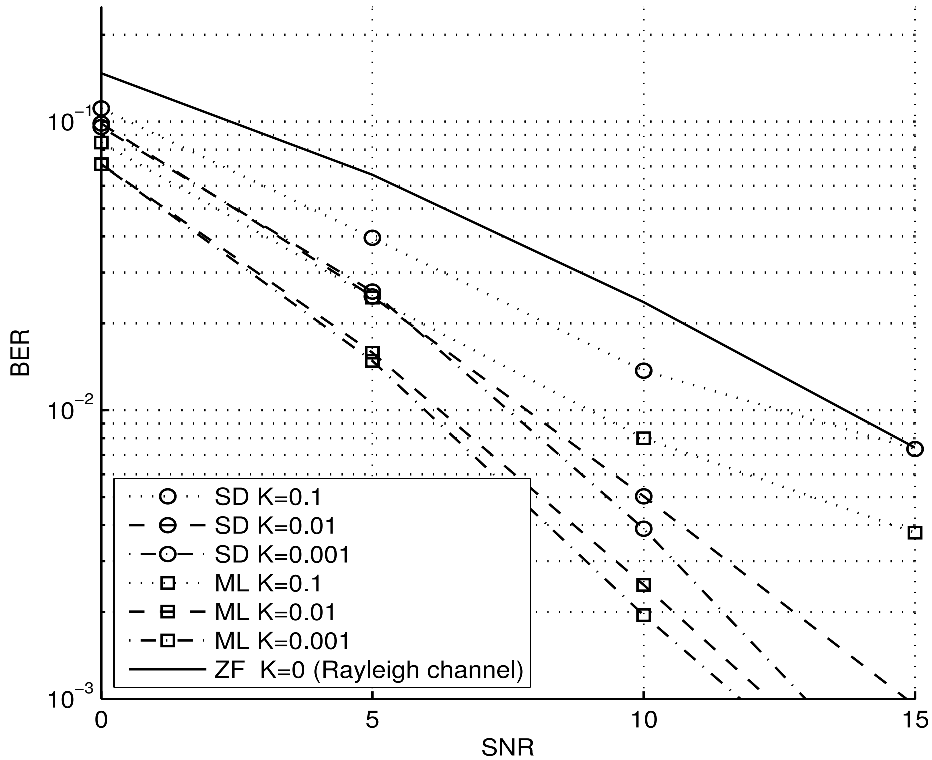

Fig. 3. BER performance graph, ML vs SD through a Rician fading channel with 2 Tx and 2 Rx, with BPSK modulation.

the other hand, SD grows far slower than ML saving a big amount of operations and making it practical to deploy.

5 Conclusions

In this paper we have studied the advantages of the spheric detector which performance is similar to ML without use as many operations making it more easy to deploy for mobility applications. Also, its lower complexity makes the SD to offer a higher throughput that the ZF detector.

Even under conditions of high correlation or fading, the proposed SD detector shows superior results in comparison with linear detectors such as ZF which its main advantage it is the low complexity but lacks of good performance. Additionally, with the rapid development of electronic devices (faster processors), the low levels of complexity and high performance, SD is set as a candidate for implementation in wireless systems of multiple antennas under critical conditions such a highly-correlated or Rician channels.

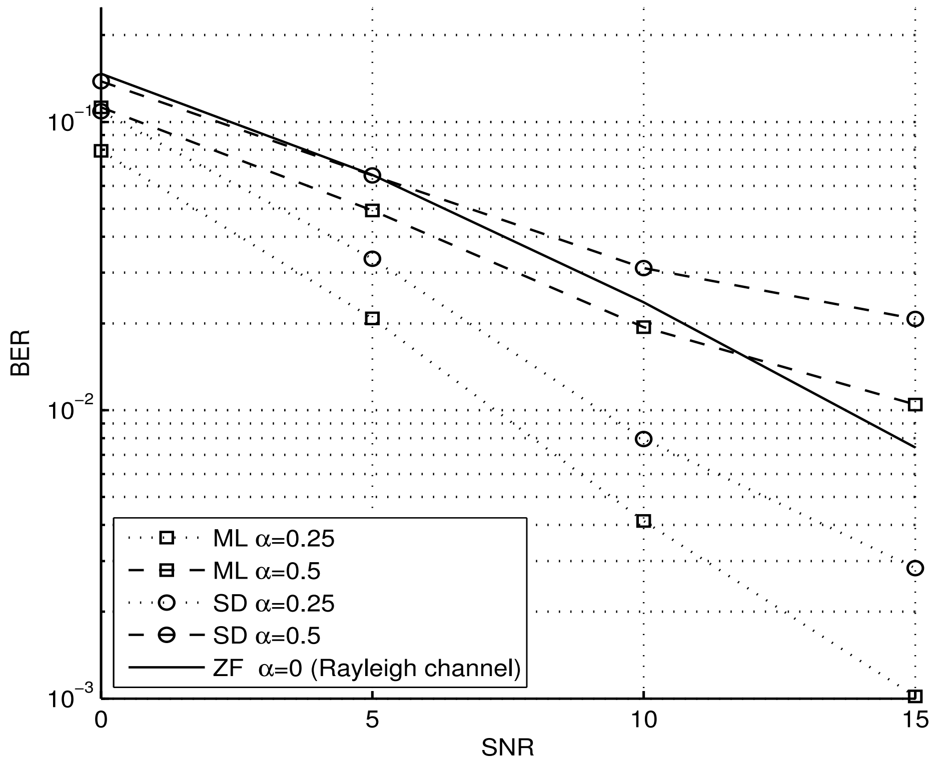

Fig. 4. BER performance graph, ML vs SD through a Correlated fading channel with 2 Tx and 2 Rx, with BPSK modulation.

Acknowledgements

This work was suppoted under grants PROMEP/103.5/10/4520, PROMEP/103.5/09/4245 and FAI: C10-FAI-05-10.38.

References

1. Theodore S. Rappaport, Wireless Communications: Principles & Practice, Prentice Hall Inc, 2002.
2. Phillip Ames, John Gabor, Wireless Communications and Computing Group, Intel Corporation, "The Evolution of Third-Generation Cellular Standards", Intel Technology Journal Q", 2000.
3. G. J. Foschini and J Gans, "On limits of wireless communications in a fading environment when using multiple antennas", Bell Labs Technical Journal, vol. 1, no. 2, Lucent Technologies, pp 41-59, Autumm 1996.
4. Telatar, I. E, "Capacity of multi-antenna Gaussian channels" ATδT Bell Labs internal report, June 1995

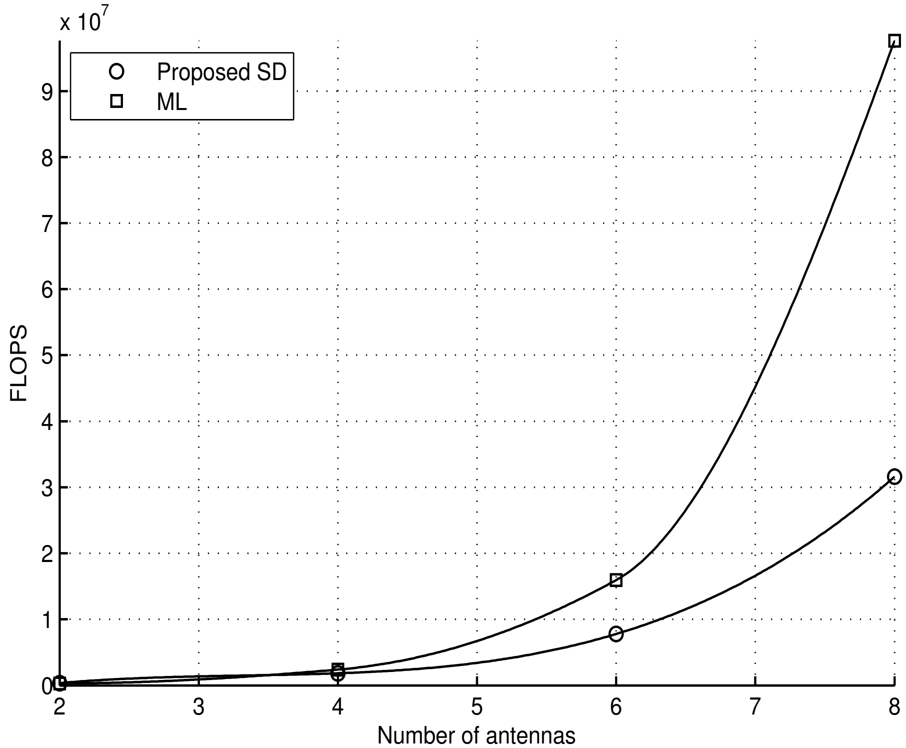

Fig. 5. Graphic of complexity, FLOPS vs. Number of antennas (Rx = Tx) and SNR = 10.

5. Joakim Jalden, Student member IEEE, "On the Complexity of Sphere Decoding in Digital Communications"
6. Oussama Damen, Ammar Chkeif, and Jean-Claude Belfiore, "Lattice Code Decoder for Space-Time Codes", IEEE communications letters, vol. 4, No. 5, May 2000.
7. B. Hassibi and H. Vikalo, "On sphere decoding algorithm I. Expected Complexity.", IEEE Transactions on Signal Processing, 53 (2005), pp. 2806-2818
8. Ulises Pineda Rico, "Link Optimisation for Mimo Communication Systems", LAP Lambert, 2010.