Efficient detection algorithms for Multiple-Input Multiple-Output (MIMO) systems

Author: Sandra Roger Varea

Directors: Alberto González Salvador
Vicenç Almenar Terré

Abstract

In the last ten years, one of the most significant technological developments that will lead to the new broadband wireless generation is the communication via Multiple-Input Multiple-Output (MIMO) systems. MIMO systems are known to provide an increase of the maximum rate, reliability and coverage of current wireless communications. Maximum-Likelihood (detection over Gaussian MIMO channels is shown to get the lowest Bit Error Rate for a given scenario. However, it has a prohibitive complexity which grows exponentially with the number of transmit antennas and the size of the constellation. Motivated by this, there is a continuous search for computationally efficient optimal or suboptimal detectors.

In this work, we carry out an state of the art review of detection algorithms and propose the combination of a suboptimal MIMO detector called K-Best Sphere Decoder with a channel matrix condition number estimator to obtain a versatile combined detector with predictable performance and suitable for hardware implementation. The effect of the channel matrix condition number in data detection is exploited in order to achieve a decoding complexity lower than the one of already proposed algorithms with similar performance. Some practical algorithms for finding the 2-norm condition number of a given channel matrix and for performing the threshold selection are also presented and their computational costs and accuracy are discussed.

Resumen

Uno de los desarrollos tecnológicos más significativos de la última década que llevarán a la nueva generación de banda ancha en movilidad es la comunicación mediante sistemas de múltiples entradas y múltiples salidas (MIMO). Los sistemas MIMO proporcionan un notable incremento en la capacidad, fiabilidad y cobertura de las comunicaciones inalámbricas actuales. Se puede demostrar que la detección óptima o de máxima verosimilitud (ML) en canales MIMO Gaussianos proporciona la mínima tasa de error de bit (BER) para un escenario dado pero tiene el inconveniente de que su complejidad crece exponencialmente con el número de antenas y el tamaño de la constelación utilizada. Por este motivo, hay una contínua búsqueda de detectores óptimos o subóptimos que sean más eficientes computacionalmente.
En este trabajo, se ha llevado a cabo una revisión del estado del arte de los principales algoritmos de detección para sistemas MIMO y se ha propuesto la combinación de un detector MIMO subóptimo conocido como *K-Best Sphere Decoder* con un estimador del número de condición de la matriz de canal, para conseguir un detector combinado basado en umbral con complejidad predecible y adecuado para implementación en *hardware*. Se ha explotado el efecto del número de condición en la detección de datos para disminuir la complejidad de los algoritmos de detección existentes sin apenas alterar sus prestaciones. Por último también se presentan distintos algoritmos prácticos para encontrar el dos número de condición así como para realizar la selección del umbral.

Author: Roger Varea, Sandra, email: sanrova@iteam.upv.es
Directors: González Salvador, Alberto, email: agonzal@dcom.upv.es
Almenar Terré, Vicenç, email: valmenar@dcom.upv.es
Submitting Date: 15-09-08

Contents

Chapter 1

Introduction

1.1 MIMO overview

In radio communications, Multiple-Input and Multiple-Output, or MIMO, refers to the use of multiple antennas at both the transmitter and receiver sides to improve communication performance. It is one of several forms of smart antenna (SA), and the state of the art of SA technology.

MIMO technology has attracted attention in wireless communications, since it offers significant increases in data throughput and link range without additional bandwidth or transmit power. It achieves this by higher spectral efficiency (more bits per second per Hertz of bandwidth) and link reliability or diversity (reduced fading). Because of these properties, MIMO is a current field of international wireless research [11].

MIMO can be sub-divided into three main categories, precoding, spatial multiplexing, or SM, and diversity coding.

- **Precoding** is multi-layer beamforming in a narrow sense or all spatial processing at the transmitter in a wide-sense. In (single-layer) beamforming, the same signal is emitted from each of the transmit antennas with appropriate phase (and sometimes gain) weighting such that the signal power is maximized at the receiver input. The benefits of beamforming are to increase the signal gain from constructive combining and to reduce the multipath fading effect. In the absence of scattering, beamforming results in a well defined directional pattern, but in typical cellular conventional beams are not a good analogy. When the receiver has multiple antennas, the transmit beamforming cannot simultaneously maximize the signal level at all of the receive antenna and precoding is used. Note that precoding requires knowledge of the channel state information (CSI) at the transmitter.

- **Spatial multiplexing** requires MIMO antenna configuration. In spatial multiplexing, a high rate signal is split into multiple lower rate streams and each stream is transmitted from a different transmit antenna in the same frequency channel. If these signals arrive at the receiver antenna array with sufficiently different spatial signatures, the receiver can separate these streams, creating parallel channels for free. Spatial multiplexing is a very powerful technique for increasing channel capacity at higher Signal to Noise Ratio (SNR). The maximum number of spatial streams is limited by the lesser in the number of antennas at the transmitter or receiver. Spatial multiplexing can be used with or without transmit channel knowledge.

- **Diversity coding** techniques are used when there is no channel knowledge at the transmitter. In diversity methods a single stream (unlike multiple streams in spatial multiplexing) is transmitted, but the signal is coded using techniques called space-time coding. The signal is emitted from each of the transmit antennas using certain principles of full or near orthogonal coding. Diversity exploits the independent fading in the multiple antenna links to enhance signal diversity. Because there is no channel knowledge, there is no beamforming or array gain from diversity coding.

Spatial multiplexing can also be combined with precoding when the channel is known at the transmitter or combined with diversity coding when decoding reliability is in trade-off. Spatial multiplexing techniques makes the receivers very complex, and therefore it is typically combined with Orthogonal frequency-division multiplexing (OFDM) or with Orthogonal Frequency Division Multiple Access (OFDMA) modulation, where the problems created by multi-path channel are handled efficiently. The IEEE 802.16e standard incorporates MIMO-OFDMA. The IEEE 802.11n standard, which is expected to be finalized soon, recommends MIMO-OFDM. MIMO is also planned to be used in Mobile radio telephone standards such as recent 3GPP and 3GPP2 standards. In 3GPP, High-Speed Packet Access plus (HSPA+) and Long Term Evolution (LTE) standards take MIMO into account. Moreover, to fully support cellular environments MIMO research consortia including IST-MASCOT propose to develop advanced MIMO techniques, i.e., multi-user MIMO (MU-MIMO).

1.2 The detection problem

Maximum-likelihood (ML) detection over Gaussian Multiple-Input Multiple-Output (MIMO) channels is shown to get the lowest Bit Error Rate (BER) for a given scenario [3]. However, it has a prohibitive complexity which grows exponentially with the number of transmit antennas and the size of the constellation. Motivated by this, there is a continuous search for computationally efficient detectors, as the well-known suboptimal linear detectors based on the ZF or MMSE approaches [3] or the recently proposed Sphere Decoding (SD) techniques ([10],[5]). Unfortunately, the detection throughput of regular SD algorithms is non-fixed, that can make these methods non useful for real time detection and hardware implementation. On the other hand, the SD method called K-Best [9] exhibits fixed complexity, but it has the drawback of not reaching the ML solution in all cases.

1.3 Outline of the Master Thesis

This master thesis is structured as follows.

In Chapter 2, the MIMO system model considered in our work is described. Next, the data detection problem for this system and several linear detection algorithms are presented.

Chapter 3 is dedicated to the description of tree-search detection algorithms, also commonly known as Sphere Decoders, which achieve full or almost ML performance with lower complexity than the exhaustive search. Special attention will be paid to K-Best Sphere Decoding algorithm, since it will be the base algorithm for the proposed combined detector.

Chapter 4 deals with the main contribution of this Master Thesis. Throughout this chapter, the structure of the combined K-Best Sphere Decoder proposed by the authors is detailed.

In Chapter 5 the problem of finding channel matrix condition number estimators is overcome and Chapter 6 focuses on proposing some threshold selection methods.

Chapter 7 denotes to the mixed problem of channel estimation and data detection, in particular, the application of Sphere Decoding methods for this purpose has been carried out.

Results appear in Chapter 8 and a summary and conclusions are drawn in Chapter 9.

Chapter 2

System model and detection algorithms for MIMO

Throughout this chapter the data detection problem in MIMO will be stated and several MIMO detection algorithms will be described.

2.1 System model

Present work is focused on the well-known Bell-Labs Layered Space Time system (BLAST) [6], illustrated by Fig. 2.1, although its application is not limited to this particular case. BLAST is a high speed wireless communication system that employs multiple antennas at both the transmitter and the receiver. In a BLAST system, the data stream is split equally into n_T transmit antennas and simultaneously sent to the channel thus overlapping time and frequency. The signals are received by n_R receive antennas as shown in Fig. 2.1 and the receiver has the task of separating the received signals in order to recover the transmitted data.

Let us consider a BLAST MIMO system characterized as block fading (the channel remains constant along the whole transmission of a data block), with n_T transmit antennas, n_R receive antennas, $n_R \geq n_T$, and a signal to noise ratio denoted by ρ. The baseband equivalent model for such MIMO system is given by

$$\mathbf{x_c} = \mathbf{H_c}\mathbf{s_c} + \mathbf{v_c}, \tag{2.1}$$

where $\mathrm{s_c}$ represents the baseband signal vector transmitted during each symbol period formed by elements chosen from the same constellation such as M-QAM. Vector $\mathbf{x_c}$ in (2.1) denotes the received symbol vector and $\mathbf{v_c}$ is a complex white Gaussian noise vector with zero mean and power N_o. Sometimes for simplicity the noise is considered with unit variance. The Rayleigh fading channel matrix $\mathbf{H_c}$ is formed by $n_R \times n_T$ complex-valued elements, h_{ij}, which represent the complex fading gain from the j-th transmit antenna to the i-th receive antenna. Moreover, the channel matrix $\mathbf{H_c}$ is considered known at the receiver for simplicity. It should be taken into account that in order to apply some detec-

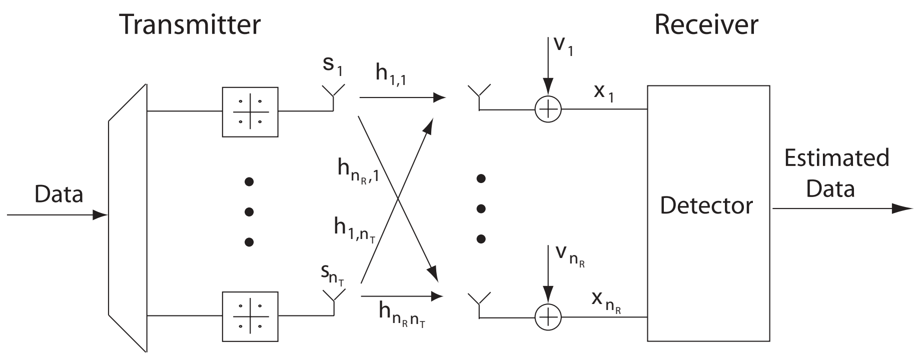

Figure 2.1: BLAST system model with n_T transmit antennas and n_R receive antennas.

tion methods to the system model (2.1), for instance the Sphere Decoding techniques, the complex model is usually transformed into a real one [9].

$$\begin{bmatrix} \Re(\mathbf{x_c}) \\ \Im(\mathbf{x_c}) \end{bmatrix} = \begin{bmatrix} \Re(\mathbf{H_c}) & -\Im(\mathbf{H_c}) \\ \Im(\mathbf{H_c}) & \Re(\mathbf{H_c}) \end{bmatrix} \begin{bmatrix} \Re(\mathbf{s_c}) \\ \Im(\mathbf{s_c}) \end{bmatrix} + \begin{bmatrix} \Re(\mathbf{v_c}) \\ \Im(\mathbf{v_c}) \end{bmatrix} \tag{2.2}$$

From now on, the real form of the system (2.1) will be considered, where the real equivalent system will be (2.2)[1]. The signal vectors will be next denoted as $\mathbf{x}$, s and $\mathbf{v}$, and the real channel matrix will be now called $\mathbf{H}$. Note that the dimensions of these vectors and matrix will be changed, the size of $\mathbf{x}$ and $\mathbf{v}$ will become $2n_R \times 1$, vector s will be considered a $2n_T \times 1$ vector and thus the channel matrix $\mathbf{H}$ will have $2n_R \times 2n_T$ entries.

2.2 Detection algorithms

It can be seen in Fig. 2.1 that the receive antennas see the superposition of all the transmitted signals. Given the received signal $\mathbf{x}$, the detection problem consists on determining the transmitted vector ŝ with the highest a posteriori probability. This is typically carried out in practice by means of solving the following least squares problem

$$\hat{\mathbf{s}} = \arg \min_{s \epsilon M^{2n_T}} ||\mathbf{x} - \mathbf{Hs}||^2 , \tag{2.3}$$

where $||\cdot||$ denotes the 2-norm and ŝ is an $2n_T$-dimensional vector with entries belonging to a M-ary alphabet. Eq. (2.3) is often called the Maximum Likelihood detection rule.

[1] $\Re(\cdot)$ and $\Im(\cdot)$ stand for the real and imaginary parts respectively.

Fig. 2.2 shows the classification of nearly most of the existing MIMO detection algorithms. In a first level, detection algorithms can be classified between ML (or exact) methods and almost ML methods. A second classification depends on the way of performing the detection that can be either in a linear way (just multiplying by a matrix in reception) or carrying out a successive interference cancellation (SIC) or via a tree search (Sphere Decoders).

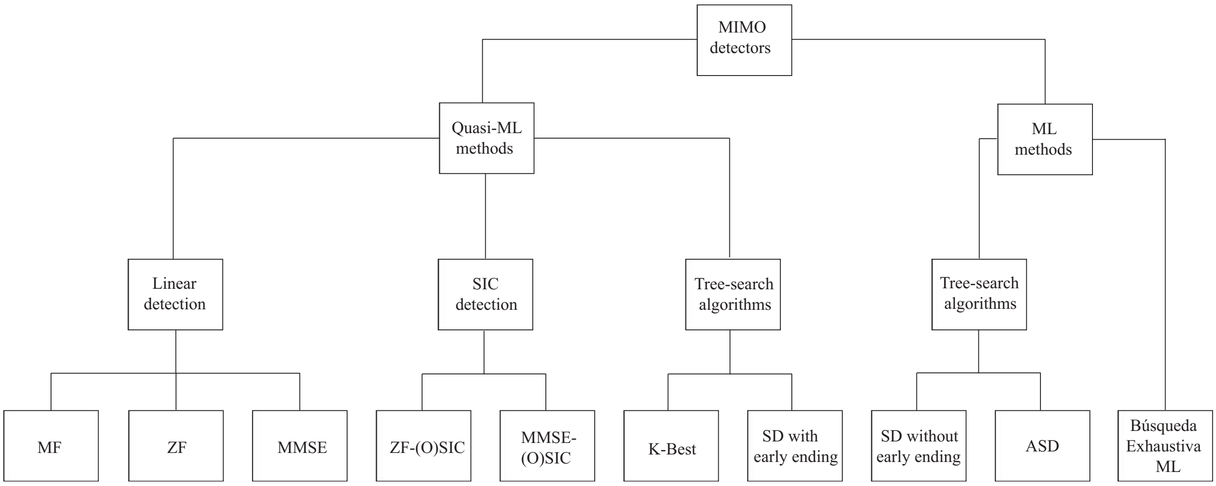

Figure 2.2: Classification of several MIMO detection algorithms.

All the BLAST detection algorithms presented above will be described throughout the rest of this chapter and the following one.

2.2.1 Maximum Likelihood (ML) detector via exhaustive search

The ML detector is the optimal detector in terms of BER since it gets the exact solution of the ML detection rule (2.3). Due to the fact that all the possible s vectors belong to a finite $2n_T$-dimensional lattice, the simplest way of finding the solution of (2.3) is performing an exhaustive search of points in the lattice and selecting the one that minimizes (2.3). This strategy leads to a very complex algorithm, with a computational cost exponentially growing with the number of transmit antennas and the size of the constellation. Alternative detectors have been developed in order to decrease this high cost in spite of losing performance and the most important ones will be in what follows presented.

2.2.2 Matched Filter (MF) detector.

The MF detector appeared as an extension of the data detection in SISO (Single-Input Single-Output) channels [2]. The detection step is carried out just by multiplying the received vector by the transpose and conjugate of the channel matrix[2]

$$\hat{\mathbf{s}} = \mathbf{quantize}\{\mathbf{H^H x}\}. \tag{2.4}$$

This algorithm exhibits near optimum behavior when the columns of $\mathbf{H}$ are close to be orthogonal, since it means that the several channels that exist in parallel are almost independent among themselves.

2.2.3 Zero Forcing (ZF) detector.

The ZF detector considers the signal from each transmit antenna as the target signal and the rest of signals as interferers [2]. The main goal of this detector is setting the interferers amplitude to zero and this is done by inverting the channel response and rounding the result to the closest symbol in the alphabet considered. When the MIMO channel matrix is square ($n_R = n_T$) and non-singular (invertible) the inversion step is performed just using the inverse of the channel matrix

$$\hat{\mathbf{s}} = \mathbf{quantize}\{\mathbf{H^{-1} x}\}. \tag{2.5}$$

However, when the channel matrix is tall ($n_R > n_T$), the pseudo inverse of $\mathbf{H}$ is then used, what leads to the following inversion step

$$\hat{\mathbf{s}} = \mathbf{quantize}\{(\mathbf{H^H H})^{\mathbf{-1}}\mathbf{H^H x}\}. \tag{2.6}$$

The ZF detector presents the problem of, in some cases, finding singular channel matrices that are not invertible. Another disadvantage is the fact that ZF focuses on cancelling completely the interference at the expense of enhancing the noise [2]. Motivated by this, the MMSE detector appeared.

2.2.4 Minimum Mean Square Error (MMSE) detector

The MMSE detector [2] minimizes the error due to the noise and the interference combined. This is done by using the following detection step:

$$\hat{\mathbf{s}} = \mathbf{quantize}\{(\mathbf{H^H H} + N_o\mathbf{I})^{\mathbf{-1}}\mathbf{H^H x}\}, \tag{2.7}$$

where $\mathbf{I}$ denotes an identity matrix.

[2] Also, a quantization step is needed to round off the result to the closest symbol in the alphabet considered, here it comes represented by the function **quantize**.

2.2.5 Nulling and cancellation detectors

The performance of already presented algorithms (ZF and MMSE) can be improved by using nonlinear techniques as symbol cancellation [3]. By using symbol cancellation, an already detected and quantized symbol from each transmit antenna is extracted out from the received signal vector, similarly to what is done in decision feedback equalization or multiuser detection with successive interference cancellation (SIC). Therefore, as soon as a signal is detected, the next one will see one interferer less.

Nulling and cancellation detectors with ordering

Nulling and cancellation detectors have the drawback of adding interference to the next symbols to be detected, when there has been any wrong decision in the already detected symbols. It can be shown that it is advantageous to find and detect first the symbols with the highest signal to noise ratio, i.e., the most reliable ones. This strategy is known as nulling and cancellation with ordering (O-SIC) [12].

Unfortunately, none of the already presented linear and SIC algorithms can reach the ML solution in all cases. This drawback will be overcome in the next chapter by employing SD algorithms.

Chapter 3

Sphere Decoding Algorithms

In this chapter, Sphere Decoding (SD) algorithms will be described and their advantageous behavior will be also discussed. Sphere Decoding (SD) methods have the ability of reaching the ML solution at lower complexity than the exhaustive search, by looking for the ML solution just within a hypersphere centered at the received signal vector. This will be shown graphically and mathematically throughout this chapter.

3.1 Sphere Decoding Fundamentals

The main interest of Sphere Decoding methods is that instead of performing an exhaustive search over the total $2n_T$-dimensional lattice points, these methods [10] limit the search for the solution to only the lattice points located within a distance of the received vector lower than a given maximum distance, called sphere radius R. The sphere radius constraint can be included in the in the ML detection rule as follows

$$\hat{\mathbf{s}} = \arg \min_{s \epsilon M^{2n_T}} \{||\mathbf{x} - \mathbf{Hs}||^2 \leq R\}, \tag{3.1}$$

For instance, Fig. 3.1 shows the lattice points of a 2x2 MIMO system using a BPSK constellation. It can be seen that if a sphere radius R is chosen, there are two lattice points that lie inside the sphere, these two points represent the candidate solutions. The ML solution would then be the closest lattice point of the list of candidate points, which is labelled in the figure as *ML*. These methods can substantially reduce the decoding complexity, however, it is necessary to find a suitable value of the sphere radius, what can be difficult in practice. The importance of the sphere radius will be discussed below.

A QR factorization of the channel matrix ($\mathbf{H} = \mathbf{QR}$) allows transforming the system (2.3) of Chapter 2 to an equivalent one that can be solved using a tree structure [10]. Matrix $\mathbf{Q}$ is orthogonal, $\mathbf{QQ^T} = \mathbf{I}$, and matrix $\mathbf{R}$ can be decomposed into an upper triangular $2n_T \times 2n_T$ matrix, denoted by $\mathbf{R}_1$, and a $(2n_R - 2n_T) \times 2n_T$ matrix of zeroes. In case

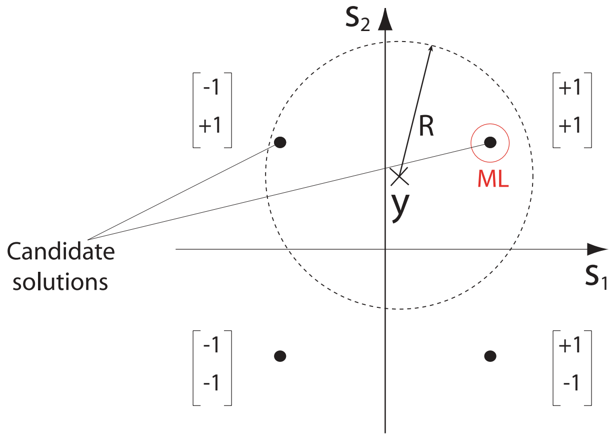

Figure 3.1: Decoding sphere of radius R for limiting the candidate lattice points in a 2x2 MIMO system using a BPSK constellation.

of multiplying (2.3) by $\mathbf{Q^T}$ and calling $\mathbf{y} = \mathbf{Q^T x}$, the problem (2.3) can be equivalently expressed as

$$\hat{\mathbf{s}} = \arg \min_{s \epsilon M^{2n_T}} \{||\mathbf{y} - \mathbf{R_1 s}||^2 \leq R\}, \tag{3.2}$$

or in a more detailed way as

$$\hat{\mathbf{s}} = \arg \min_{s \epsilon M^{2n_T}} \{\sum_{i=1}^{2n_T} |y_i - \sum_{j=i}^{2n_T} r_{ij} s_j|^2 \leq R\}. \tag{3.3}$$

where the triangular structure of $\mathbf{R_1}$ has also been exploited.

In order to solve (3.3) via a tree search, the following recursion is performed for $i = 2n_T, \dots, 1$:

$$T_i(S^{(i)}) = T_{i+1}(S^{(i+1)}) + |e_i(S^{(i)})|^2 \tag{3.4}$$

$$e_i(S^{(i)}) = y_i - \sum_{j=i}^{2n_T} r_{ij} s_j, \tag{3.5}$$

where i denotes each tree level, $S^{(i)} = [s_i, s_{i+1}, \dots, s_{2n_T}]$, $T_i(S^{(i)})$ is the accumulated Partial Euclidean Distance (PED) up to level i, where $T_{2n_T+1}(S^{(2n_T+1)}) = 0$, and $|e_i(S^{(i)})|^2$ is the distance between levels i and $i+1$ in the decoding tree, which will be represented as the branch weight. Partial solutions are represented as nodes n and nodes are expanded in order to look for the ML solution or the closest lattice point. It is required to find the ML solution expanding as few nodes as possible in order to reduce the computational effort. Fig. 3.2 depicts the decoding tree associated to the decoding sphere of Fig. 3.1, note that the tree will have as many levels as transmit antennas in the system (for complex constellations this number of levels will be doubled) and each node will have as many children nodes as the constellation size (for complex constellations the size of the equivalent real constellation will be considered instead). It can be seen that the search for the solution is performed in two levels, in each level branches with an accumulated PED higher than the sphere radius are discarded, resulting in less visited points.

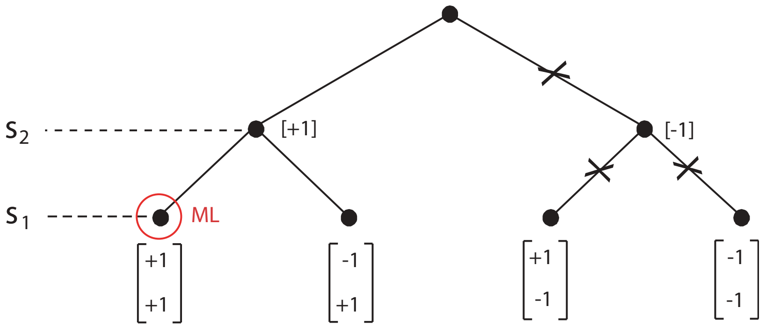

Figure 3.2: Decoding tree associated to the decoding sphere of Fig. 3.1.

Different tree search strategies have been proposed, some of them can be found in [9],[10] and [15] but they can be classified into two main types of tree search: Depth-First and Breadth-First. In the Depth-First algorithms the tree is explored beginning from the root descending to the leaf nodes, but exploring every child node from left to right. Fig. 3.3 makes clear this kind of search.

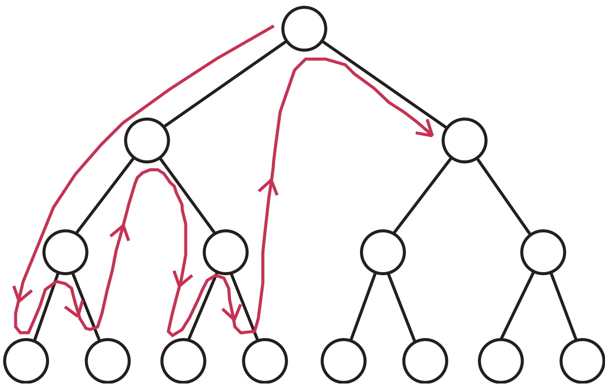

Figure 3.3: Decoding tree where a Depth-First strategy is followed.

In the Breadth-First algorithms the tree is explored descending level by level up to the leaf nodes, every child in the same level has to be visited before starting to visit the following level. Fig. 3.4 depicts the general idea behind Breadth-First algorithms.

As said above, a suitable sphere radius is generally needed for getting the ML solution expanding as few nodes as possible. However, if a too small sphere radius is chosen, there can be no candidate solutions and the algorithm will not perform correctly. On the other hand, if a too large sphere radius is selected, too many candidate points may

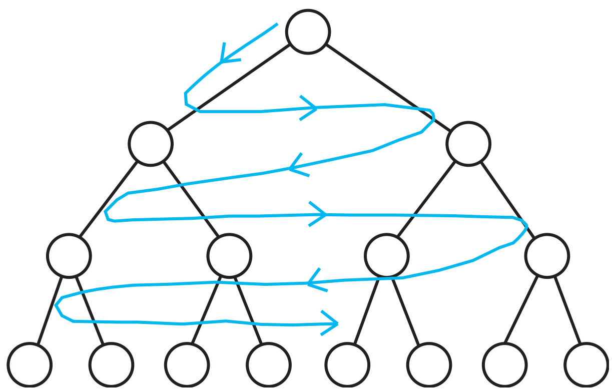

Figure 3.4: Decoding tree where a Breadth-First strategy is followed.

be found and the complexity of the algorithm can equal the one of an ML exhaustive search, without any advantage over existing methods. Fig. 3.5 shows the candidate points inside two spheres of different radius, for a 2D case. Selecting the smaller radius (R_1) provides just one candidate point and choosing R_2 leads to a search among four candidate points. There exist several estimates of the radius [10], for instance the ones that set the sphere radius as the distance between the received vector and the solution provided by a low complexity detection method as ZF or MMSE. Other authors suggest choosing a scaled version of the noise variance as a candidate radius, since it seems obvious that the transmitted vector will be moved away from its original position a distance related to the variance of the noise in the system.

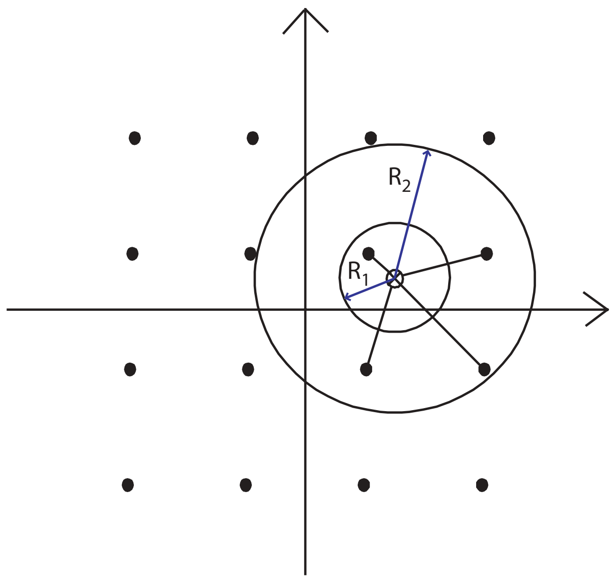

Figure 3.5: Comparison between the number of candidate points inside spheres of radius R_1 and R_2.

3.2 Fincke-Pohst and Schnorr-Euchner Sphere Decoders

One of the earliest Depth-First Sphere Decoding algorithms that appeared is the one that follows the Fincke-Pohst (FP-SD) strategy [5]. The main feature of this algorithm is that candidate solutions are discarded based on a sphere radius that has to be selected in a correct way.

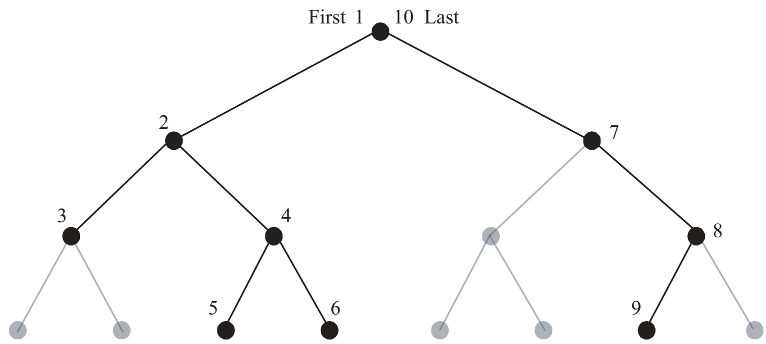

Figure 3.6: Decoding tree for a 3x3 MIMO system with a BPSK constellation which follows a Fincke-Pohst search strategy.

Let us show how this method works with an example. It can be seen in Fig. 3.6 a decoding tree for a 3x3 BPSK MIMO system. It will be considered that an appropriate sphere radius has been selected before starting the search for the solution. The order in which nodes are visited is denoted by ascending numbers. Branches and nodes in gray color correspond to solutions that have been discarded because their PEDs are known to exceed the sphere radius. It can be seen that after traversing the three levels of the tree, there are just three possible solutions inside the sphere, with leaf nodes numbered as 5, 6 and 9. The final solution would then be the path with the minimum sum of PEDs from the root among this three candidates.

The Schnorr-Euchner SD (SE-SD) also performs a Depth-First search but it refines the FP algorithm for computing even less number of nodes. Instead of exploring the tree from left to right, the SE-SD computes the PEDs of the children nodes from a given node in the current level and explores them in increasing order of their branch weights. This improvement leads to reaching valid leaf nodes quickly than by using the FP-SD. Unfortunately, the number of branch and node weights to compute remains the same. For overcoming this problem, SE-SD proposes changing the search radius adaptively once a leaf node has been discovered, since after having discovered a point in the search set, we become interested only in locating those points that are even closer to the target than that point.

3.3 K-Best Sphere Decoder

K-Best Sphere Decoder [9] is a Breadth-First algorithm that expands only those K survivor nodes that show the smallest accumulated PEDs at each level of the decoding tree, see Fig. 3.7.

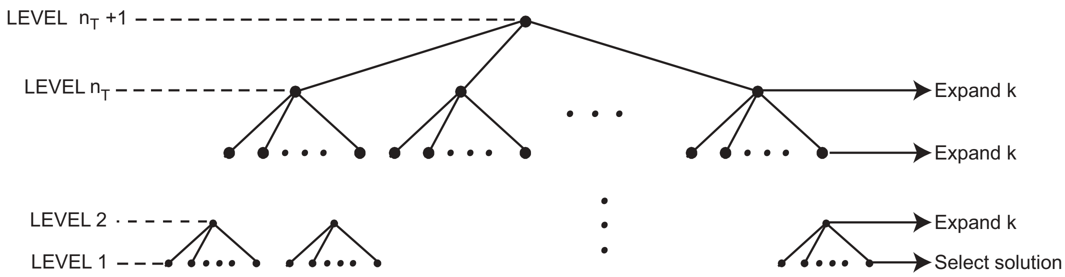

Figure 3.7: Decoding tree of the K-Best algorithm.

This method has a considerable difference respect to Fincke-Pohst, since now the candidate solutions are not discarded using a sphere radius but having a list with a fixed number of closest lattice points (K) up to the current level. The detected signal vector $\hat{\mathbf{s}}$ is given by the path from the root up to the leaf node with the smallest total Euclidean distance.

The main advantage of this method is that the maximum number of paths is limited, that yields a fixed computational effort and makes the algorithm hardware implementation easier. Variants of this algorithm also include a sphere radius in order to reduce the number of explored paths [13] but unfortunately, this number is then non-fixed and unknown.

As it is shown in [13], it is more likely to discard the ML solution at early decoding stages, since in latest levels the accumulated PED is closer to the final total distance. Thus, the method can also be modified to work with different K values at different decoding levels, which is called Dynamic K-Best. Dynamic K-Best will have the disadvantage of not having the same complexity at every level.

3.4 Automatic Sphere Decoder

The Automatic Sphere Decoder (ASD) was initially proposed in [15]. It is a Breadth-First algorithm that does not make use of a sphere radius to find the solution. It stores a list that defines the limit between the already explored part of the tree and the non-explored yet. At the beginning of the algorithm, the list only contains the root node and its associated accumulated distance, which is equal to zero. In each iteration, the ASD selects and expands the node inside the list that has the smallest associated distance. This just expanded node, is removed from the list and replaced by its children nodes. The first time that a leaf node is selected for expansion the detector returns the associated candidate solution and ends.

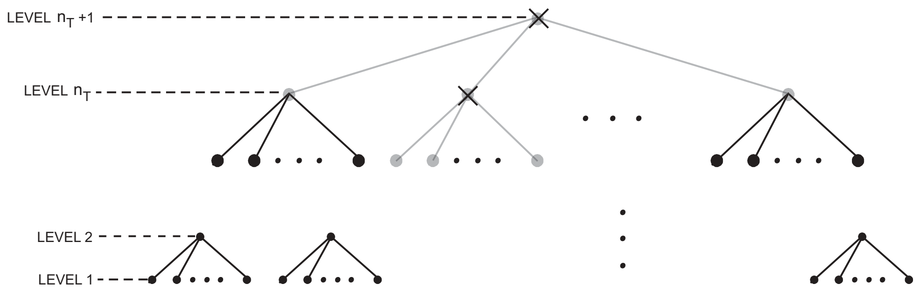

Figure 3.8: Decoding tree of the ASD algorithm.

In Fig. 3.8 it can be seen the decoding tree when just the root node and a node located at the n_T level have been expanded. The already visited branches are depicted in gray.

The main disadvantage of this method is the need for a variable size list of nodes, what makes its Hardware implementation more difficult.

Chapter 4

Combined K-Best Sphere Decoder

Experiments [1] show that the channel matrix condition number is strongly related to the performance of suboptimal detection schemes, since it is a measure of how the original constellation is distorted by the channel. For instance, Fig. 4.1 shows the performance degradation of some MIMO detectors with the increase of the channel matrix condition number for a value of signal to noise ratio $\rho = 20dB$. Note that it can be found a channel matrix condition number value that makes the performance of 5-Best equal to the 8-Best or even to the ML. Also, as the condition number gets higher, the performance of 8-Best is not as much degraded as the one of 5-Best. Considering this, it seems obvious that a suitable combination of both algorithms based on the channel condition number could have an almost ML behavior for this value of ρ. It can also be shown that the condition number increases with the size of the channel matrix [7], so for a higher number of antennas, the detection degradation will increase.

Authors of [14] have developed combined detectors based on condition number thresholding, for instance ML and ZF, but so far their complexities are generally non-fixed. Furthermore, another disadvantage of these algorithms is the need for implementing two or more different decoders to build the combined one. Therefore, we suggest to use a unique algorithm, the K-Best SD, and just change subsequently its parameter K between a set of acceptable values previously selected, which is obviously more suitable for real systems. Moreover, it will be necessary to develope some condition number estimator for determining which detector is the most appropriate for each channel. In Chapter 5 a low complexity 2-norm condition number estimator that makes use of the QR factorization [8], which is always available when working with SD methods [10], will be developed. This estimator together with the Power Method for computing eigenvalues [8] can provide a useful approximation of the condition number of the channel matrix. As a result, the combined decoder including the condition number estimator will have an adjustable and predictable complexity, so it will become suitable for practical systems.

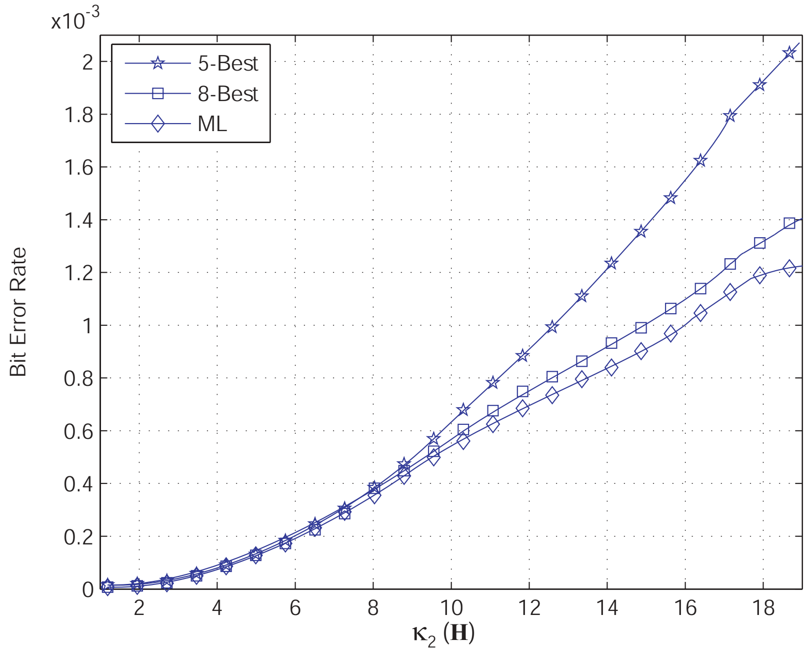

Figure 4.1: BER of different MIMO detectors with a 16-QAM constellation, a 4x4 MIMO channel and a value of $\rho = 20dB$, as a function of the channel matrix condition number $\kappa_2(\mathbf{H})$.

4.1 Combined decoder

The proposed combined decoder always works with a K-Best detector but it can select a low value of K while working with well-conditioned channels and switch to a higher value of K whether the channel is unfortunately ill-conditioned. This way, a greater decoding tree is visited when dealing with poor channels, thus there is less probability of discarding the ML solution too early. As said before, an estimator of the condition number together with a threshold condition number, denoted by κ_{th}, are used to classify the channels and consequently adjust the K parameter.

Fig. 4.2 depicts the flow diagram of this combined sphere decoder and the steps to perform the combined detection are next detailed. Firstly a threshold condition number is chosen. It must be checked whether $\mathbf{H}$ changed, in case it did not change, the currently fixed K-Best is used, otherwise the new channel condition number has to be estimated. Then, using the threshold κ_{th}, the suitable K value is selected between k_1 and k_2 with $k_2 > k_1$. Finally, K-Best Sphere Decoding is used for carrying out the detection. It can be observed in Fig. 4.2 that, as said before, an estimator of the condition number together with a threshold condition number, denoted by κ_{th}, have been used to classify the channels and consequently fix the K value.

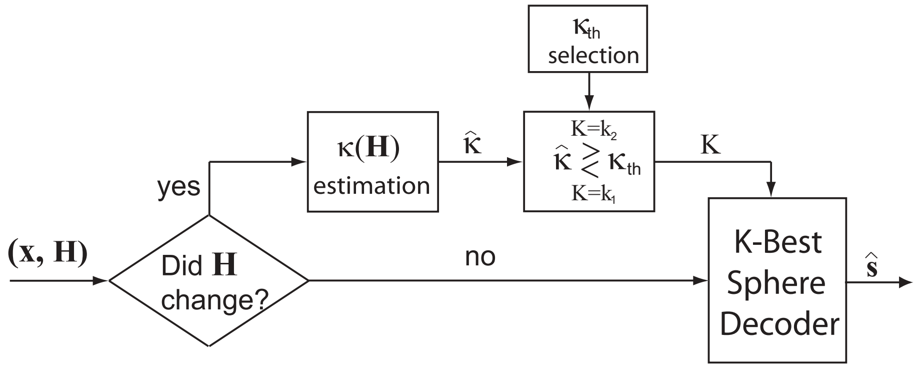

Figure 4.2: Flow diagram of a combined K-Best Sphere Decoder with threshold selection and channel condition number estimation.

The performance achieved by this combined detector will be discussed in Chapter 8.

Chapter 5

Channel matrix condition number estimation

The sensitivity of the solution of a non-singular system of linear equations $\mathbf{Ax} = \mathbf{b}$ with respect to perturbations of the matrix $\mathbf{A}$ is directly related to its condition number [4]. Although the matrix condition number depends on the selected norm, if the matrix is well-conditioned, the condition number will be small in all norms, otherwise it will be large. Thus, the most convenient norm is usually selected between the 1-norm, 2-norm and ∞-norm. The 2-norm condition number is defined as

$$\kappa_2(\mathbf{A}) = \frac{\sigma_{max}}{\sigma_{min}}, \tag{5.1}$$

being σ_{max} and σ_{min} the maximum and minimum singular values of $\mathbf{A}$ respectively. When $\mathbf{A}$ is square $n \times n$, $\kappa_2(\mathbf{A})$ can be also computed as $\|\mathbf{A}\|\|\mathbf{A^{-1}}\|$. Although other condition numbers can be considered, $\kappa_2(\mathbf{A})$ will be selected in our work because of some special properties presented below. Fig. 5.1 shows the probability density function of the 2-norm condition number for a 8x8 real Gaussian MIMO channel matrix, it can be seen that although lower condition number values are more likely, there are also channel occurrences with high condition number where the performance of suboptimal detectors is decreased, as it was introduced in Chapter 4.

5.1 Condition number estimator

In order to carry out the combined detection in practice, it is important to have a reliable estimate of $\kappa_2(\mathbf{H})$. As explained above, K-Best requires a factorization of the form $\mathbf{H} = \mathbf{QR}$. Only in case of working with 2-norm it is true that

$$\kappa_2(\mathbf{H}) = \kappa_2(\mathbf{R_1}) = \|\mathbf{R_1}\|\|\mathbf{R_1^{-1}}\|, \tag{5.2}$$

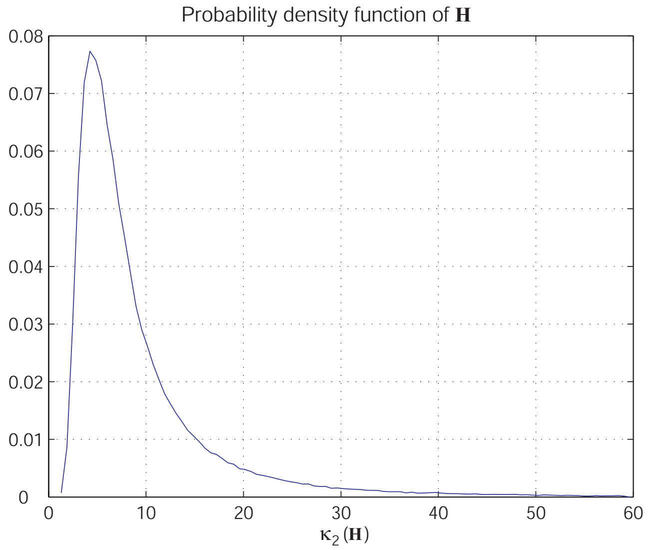

Figure 5.1: Probability density function of the 2-norm condition number of 8x8 real Gaussian channel matrices.

since $\|\mathbf{Q}\| = 1$.

Due to the fact that $\mathbf{R_1}$ is triangular, it can be noted that $\|\mathbf{R_1}\|$ can be calculated faster than $\|\mathbf{H}\|$. Moreover, $\|\mathbf{R_1}\| = \sigma_{max}$ can be efficiently computed by applying the Power Method [8] and a low complexity estimator will be proposed for calculating $\|\mathbf{R_1^{-1}}\| = 1/\sigma_{min}$.

5.1.1 The Power Method for computing $\|\mathbf{R_1}\|$

The Power Method is an iterative algorithm that obtains the largest eigenvalue of a given matrix [8]. Given a $n \times n$ diagonalizable matrix $\mathbf{A}$ with $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$ eigenvalues, this method starts with a unit 2-norm vector $\mathbf{q}^{(0)} \in \mathcal{R}^n$ as an initial approximation of one of the dominant eigenvectors. At each iteration i, it computes the new $\mathbf{q}^{(i)}$ in two steps. First, the vector $\mathbf{z}^{(i)} = \mathbf{A}\mathbf{q}^{(i-1)}$ is calculated and next it is normalized resulting in $\mathbf{q}^{(i)} = \mathbf{z}^{(i)}/\|\mathbf{z}^{(i)}\|$. After the last iteration of the process, the maximum eigenvalue can be computed as

$$\lambda_{max} = [\mathbf{q}^{(i)}]^T \mathbf{A} \mathbf{q}^{(i)}. \tag{5.3}$$

In our work, the Power Method is proposed for calculating $\|\mathbf{R_1}\| = \sigma_{max}$, previously calculating the maximum eigenvalue of $\mathbf{A} = \mathbf{R_1^T R_1}$, which corresponds to σ^2_{max} (obviously, the method can be applied to $\mathbf{R_1^T R_1}$ without computing explicitly this product). Considering that the maximum size of the channel matrices is usually up to 8x8, a maximum number of 10 iterations for running the Power Method gets quite accurate results. Thus, the number of flops can be computed as $21n^2 + 22n$.

Table 5.1: Measured Complexity in Number of Flops

Real channel	Proposed estimator for $\kappa_2(\mathbf{R_1})$	Power Method for $\kappa_2(\mathbf{R_1})$
4x4	482	850
8x8	1698	3042
16x16	6338	11458

5.1.2 Estimator of $\|\mathbf{R_1^{-1}}\|$

This method was firstly developed in [4]. Two triangular systems need to be formulated. The first one is $\mathbf{R_1^T x} = \mathbf{b}$ and $\mathbf{b}$ has to be chosen so that its solution $\hat{\mathbf{x}}$ will make $\|\hat{\mathbf{x}}\|/\|\mathbf{b}\|$ as large as possible. This is achieved by an iterative process with n steps, considering the size of $\mathbf{R_1^{-1}}$ is n. At each step i, b_i is chosen between $+1$ and -1, in order to maximize x_i, which will be computed as

$$r_{ii}x_i = b_i - (r_{1i}x_i + \ldots + r_{i-1i}x_{i-1}). \tag{5.4}$$

The second system to solve is $\mathbf{R}_1\mathbf{y} = \hat{\mathbf{x}}$. Once its solution $\hat{\mathbf{y}}$ is obtained, the estimation for $\|\mathbf{R_1^{-1}}\| = 1/\sigma_{min}$ is given by $\|\hat{\mathbf{y}}\|/\|\hat{\mathbf{x}}\|$. The computational cost of this estimator is $2n^2 + 6n$ flops.

As soon as σ_{max} and $1/\sigma_{min}$ are available, $\kappa_2(\mathbf{R_1})$ is calculated as the product of both values.

In [14] the authors estimate both σ_{max} and $1/\sigma_{min}$ by means of the Power Method, what leads to a total number of flops of $42n^2 + 44n + 2$. On the other hand, the mixed estimator that we propose has been selected in the present work mainly because it better exploits the QR factorization used in the K-Best algorithms, also it only requires $23n^2 + 28n + 2$ flops.

Fig. 5.2(a) shows that the relative error of our proposed estimator for $\kappa_2(\mathbf{R_1})$ is higher than the one of the Power Method for computing the whole condition number. However, as can be seen in Fig. 5.2(b), the error magnitude is not very significant, making this estimator useful for combined detectors. Table 5.1.2 shows that the complexity of our proposed estimator measured in number of flops is almost the half of the complexity of the estimator that only employs the Power Method. It is possible to decrease the complexity of our proposed estimator even more by estimating $\|\mathbf{R_1}\|$ in the same way that $\|\mathbf{R_1^{-1}}\|$. However, by doing this, the accuracy of the resulting condition number estimator can be not good enough for our application.

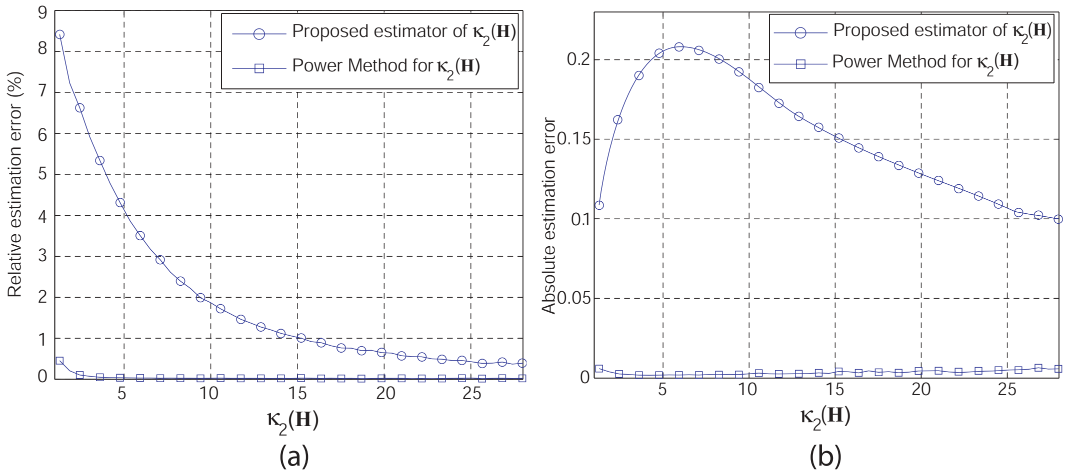

Figure 5.2: Error of our proposed estimator compared with the error of the Power Method for computing the whole condition number: (a) Relative, (b) Absolute.

Chapter 6

Threshold selection

Two ways of selecting the threshold condition number κ_{th} in a combined decoder are presented throughout this section. The first one is based on the analysis of the achieved BER of the lower performance algorithm used in the combined decoder for a given SNR. Considering the algorithm has an average BER denoted by BER_{av}, a suitable κ_{th} must guarantee that $\mathrm{BER} \leq \mathrm{BER}_{av}$ when our channel has $\kappa_2(\mathbf{H}) \leq \kappa_{th}$. This will be done in order to assure that the lower performance algorithm is only used when it has optimal performance. Taking into account these considerations, BER values of the lower performance algorithm are represented within a range of condition numbers in a similar manner than in Fig. 4.1. The abscissa value of $\kappa_2(\mathbf{H})$ for which a $\mathrm{BER} = \mathrm{BER}_{av}$ is achieved is selected as the candidate κ_{th}, since its obvious that for $\kappa_2(\mathbf{H}) > \kappa_{th}$ the achieved BER will overpass BER_{av}.

Let us clarify this threshold selection method by means of an example. Supposing that the lower performance algorithm used in the combined one is 2-Best, Fig. 6.1 depicts the achieved BER for condition numbers ranging from 1 to 30 when the value of ρ is equal to 20dB. The average BER obtained for this SNR (without condition number distinction) was $\mathrm{BER}_{av} = 1.52 \times 10^{-2}$ and this value sets the threshold to $\kappa_{th} = 9.41$. For simplicity, choosing the value of $\kappa_{th} = 10$ instead of $\kappa_{th} = 9.41$ leads also to acceptable results.

The second method for selecting the threshold states that the chosen κ_{th} should provide the desired average computational cost of the combined algorithm, which is related to the average number of expanded nodes *n*. For instance, in case of combining k_1-Best SD and k_2-Best SD, considering $k_1 < k_2$, for a threshold value of κ_{th}, the resulting average number of expanded nodes $n_{\kappa_{th}}$ would be given by

$$n_{\kappa_{th}} = (1 - P_{\kappa \geq \kappa_{th}}) n_{k_1-Best} + P_{\kappa \geq \kappa_{th}} n_{k_2-Best}, \tag{6.1}$$

where $P_{\kappa \geq \kappa_{th}}$ is the probability of having a channel with condition number higher than a threshold κ_{th} and it can be calculated as a cumulative distribution of the probability density function depicted in Fig 5.1. The number of expanded nodes in each k_i-Best al-

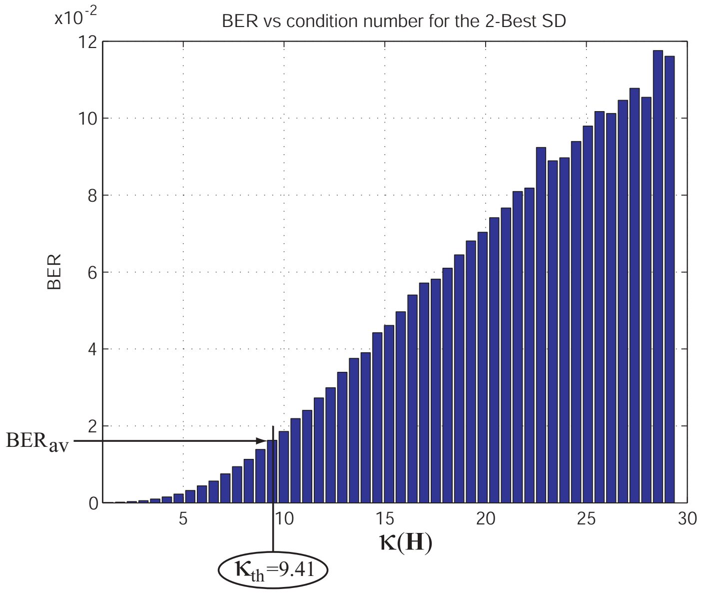

Figure 6.1: BER for the 2-Best SD with $\rho = 20$ on a 4x4 MIMO 16-QAM system as a function of the channel matrix condition number $\kappa(\mathbf{H})$.

gorithm is denoted by n_{k_i-Best} (for $i = 1, 2$). It can be noted that the desired maximum average number of expanded nodes will determine the threshold value. In the same way, for a given threshold, the number of expanded nodes will be straightforwardly predicted.

Chapter 7

Joint channel estimation and signal detection

In our system design, perfect knowledge of the channel coefficients has been assumed (perfect Channel State Information (CSI)). Note that in the introduction chapter we considered a block fading channel $\mathbf{H}$ for formulating the detection problem. In practical systems, the channel coefficients are often estimated using a training sequence, thus sacrificing a fraction of the transmission rate. The obtained channel estimate is then used for performing the typical data detection, using in most cases some of the algorithms presented in Chapters 2 and 3. This is usually called as disjoint channel estimation and data detection strategy.

In this Master Thesis the problem of Joint channel estimation and data detection has been considered. It has already been shown that carrying out both steps simultaneously can considerably reduce the computational effort without reducing the detection performance very much [16]. The computational cost decrease appears due to the fact that Sphere Decoding algorithms can be applied when the joint channel estimation and data detection problem is formulated as an integer least-squares problem. In what follows we will deal with the problem of joint channel estimation and data detection for SIMO channels. So far, no interesting results have been obtain for the MIMO case.

7.1 System model

Let us consider a SIMO channel with n_R receive antennas and constant for some time interval T. If the channel is represented by a $n_R \times 1$ channel vector $\mathbf{h}$, the received signal $\mathbf{X}$ can be written as

$$\mathbf{X} = \mathbf{h}\mathbf{s}^* + \mathbf{V} = \mathbf{h}[\mathbf{s}_\tau^* \quad \mathbf{s}_\mathbf{d}^*] + \mathbf{V}, \tag{7.1}$$

where s_τ is the $T_\tau \times 1$ vector of training symbols and s_d is the $T_d \times 1$ vector of data symbols, with $T_\tau + T_d = T$. The components of s_τ and s_d belong to a PSK constellation and they are supposed to be normalized for having energy equal to 1. Matrix $\mathbf{V}$ represents a $n_R \times$ T additive noise matrix whose elements are assumed to be independent, identically distributed (iid) complex Gaussian random variables.

7.2 The joint ML channel estimation and signal detection problem

The joint ML channel estimation and signal detection problem can be stated as follows:

$$\min_{\mathbf{h},\mathbf{s_d}} ||\mathbf{X} - \mathbf{h}\mathbf{s}^*||^2, \tag{7.2}$$

which is a mixed optimization problem: it is a least-squares problem in $\mathbf{h}$ and an integer least-squares problem in s_d. The solution to the integer least-squares problems (data detection problem) is usually found via the already presented detection techniques (see Chapters 2 and 3), but from now on we will try to solve both optimization problems simultaneously.

For any given s, the channel $\hat{\mathbf{h}}$ that minimizes (7.2) is given by

$$\hat{\mathbf{h}} = \mathbf{X}\mathrm{s}/\,||\mathrm{s}||^{\mathbf{2}} = \frac{1}{T}\mathbf{X}\mathrm{s}. \tag{7.3}$$

Substituting (7.3) into (7.2) gives

$$||\mathbf{X} - \mathbf{h}\mathbf{s}^*||^2 = \left|\left|\mathbf{X}(\mathbf{I} - \frac{1}{T}\mathbf{ss}^*)\right|\right|^2 = tr[\mathbf{X}(\mathbf{I} - \frac{1}{T}\mathbf{ss}^*)\mathbf{X}^*] = tr(\mathbf{XX}^*) - \frac{1}{T}\mathbf{s}^*\mathbf{X}^*\mathbf{Xs}, \tag{7.4}$$

and the problem (7.2) becomes equivalent to

$$\max_{\mathbf{s_d}} \mathbf{s}^*\mathbf{X}^*\mathbf{Xs}. \tag{7.5}$$

Let $\hat{\lambda} = \lambda_{max}(\mathbf{X}^*\mathbf{X})$ denote the maximum eigenvalue of $\mathbf{X}^*\mathbf{X}$, and let $\rho > \hat{\lambda}$ (for instance, it can be chosen $\rho = tr(\mathbf{X}^*\mathbf{X})$). The problem (7.5) is then equivalent to

$$\min_{\mathbf{s_d}} \mathbf{s}^*(\rho\mathbf{I} - \mathbf{X}^*\mathbf{X})\mathrm{s}. \tag{7.6}$$

For simplicity, we will call $\rho\mathbf{I} - \mathbf{X}^*\mathbf{X} = \mathbf{\Gamma}$ and $\mathbf{\Gamma}$ will be positive definite by construction. It will be useful to partition $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{11} & -\mathbf{\Gamma}_{12} \\ -\mathbf{\Gamma}_{12}^* & \mathbf{\Gamma}_{22} \end{bmatrix}, \tag{7.7}$$

where $\mathbf{\Gamma}_{11}$ is a $T_\tau \times T_\tau$ matrix, $\mathbf{\Gamma}_{12}$ is a $T_\tau \times T_d$ matrix and $\mathbf{\Gamma}_{22}$ is a $T_d \times T_d$ matrix. Since $\mathbf{\Gamma}$ is positive definite, so is $\mathbf{\Gamma}_{22}$. Therefore, it can be written

$$\begin{bmatrix} s_\tau^* & s_d^* \end{bmatrix} \begin{bmatrix} \mathbf{\Gamma}_{11} & -\mathbf{\Gamma}_{12} \\ -\mathbf{\Gamma}_{12}^* & \mathbf{\Gamma}_{22} \end{bmatrix} \begin{bmatrix} s_\tau^* \\ s_d^* \end{bmatrix} = (\mathbf{s}_d - \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau)^*\mathbf{\Gamma}_{22}(\mathbf{s}_d - \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau) + \mathbf{s}_\tau(\mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*)\mathbf{s}_\tau. \tag{7.8}$$

Since $\mathbf{s}_\tau(\mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{12}\mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*)\mathbf{s}_\tau$ does not depend on $\mathbf{s_d}$, the optimization (7.6) becomes

$$\min_{\mathbf{s}_d} (\mathbf{s}_d - \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau)^*\mathbf{\Gamma}_{22}(\mathbf{s}_d - \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau) = \min_{\mathbf{s}_d} \|\mathbf{s}_d - \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau\|^2\mathbf{\Gamma}_{22} \tag{7.9}$$

Due to the fact that $\mathbf{\Gamma}_{22}$ is positive definite, it has a Cholesky factorization of the form $\mathbf{\Gamma}_{22} = \mathbf{R}^*\mathbf{R}$ where $\mathbf{R}$ is an upper triangular matrix. If we denote $\hat{\mathbf{s}}_d = \mathbf{\Gamma}_{22}^{-1}\mathbf{\Gamma}_{12}^*\mathbf{s}_\tau$ in 7.9 and make use of the Cholesky factorization of $\mathbf{\Gamma}_{22}$, the minimization problem results in

$$\min_{\mathbf{s}_d} \|\mathbf{R}(\mathbf{s}_d - \hat{\mathbf{s}}_d)\|^2, \tag{7.10}$$

thus, it is easy to see that the problem (7.10) can be efficiently solved by means of Sphere Decoding algorithms.

Chapter 8

Results

8.1 Combined K-Best Sphere Decoder

The experimental setup used for our simulations is depicted in Fig. 8.1.

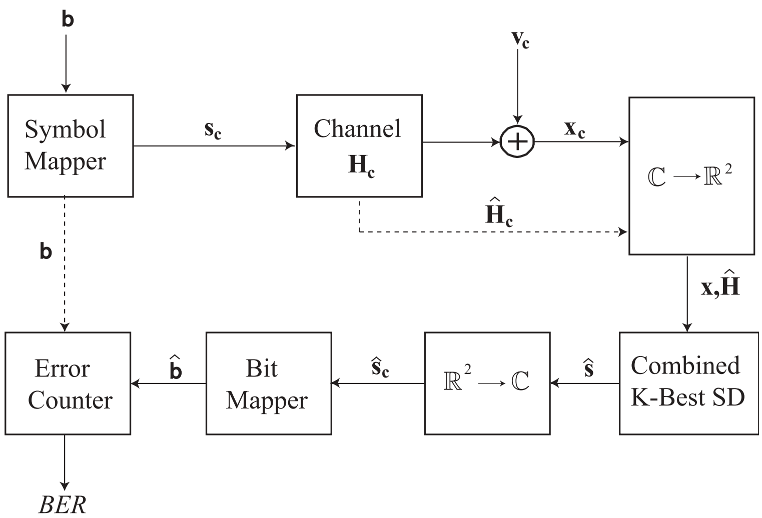

Figure 8.1: Experimental setup for testing the Combined K-Best Sphere Decoder.

The bit stream b is mapped into symbols of a constellation such as QPSK or 16-QAM forming the vector $\mathbf{s_c}$. The symbol vector is next affected by the channel and the addition of a noise vector. The received vector $\mathbf{x_c}$ needs to be transformed into its real form $\mathbf{x}$, together with the channel estimate $\hat{\mathbf{H}}_\mathbf{c}$, which is supposed to be available in the reception part. Next, the received data enters the Combined K-Best decoder for carrying out the data detection. It is supposed that the Combined decoder has an estimate of the channel condition number and also a threshold has been selected in a previous training period, this blocks are not included in the system description because they were detailed in Chapter 4. After performing the detection, the detected symbol vector s is transformed into its original complex form $\hat{\mathbf{s}}_c$. The detected symbol vector is then demapped into the vector of bits $\hat{\mathbf{b}}$. Finally both $\hat{\mathbf{b}}$ and $\mathbf{b}$ are the inputs of a Bit Error Counter, which will provide the Bit Error Rate of the algorithm as an output.

A 4x4 MIMO system and a 16-QAM alphabet were considered for our simulations. A bitstream of 10000 bits was mapped into symbols and sent in groups of 4 simultaneously by the 4 transmit antennas. The combined algorithm was formed by the 2-Best SD and the 12-Best SD. The threshold value $\kappa_{th} = 10$ was chosen following the first proposed threshold selection method, see Chapter 6 and a higher threshold value was chosen to compare results, in this case $\kappa_{th} = 20$. The algorithm was run first using the exact 2-norm channel matrix condition number and afterwards using the estimator proposed in Chapter 5. As expected, Fig. 8.2 illustrates that the performance gets worse as the threshold increases and estimating the 2-norm condition number does not change the performance of the combined detectors at all.

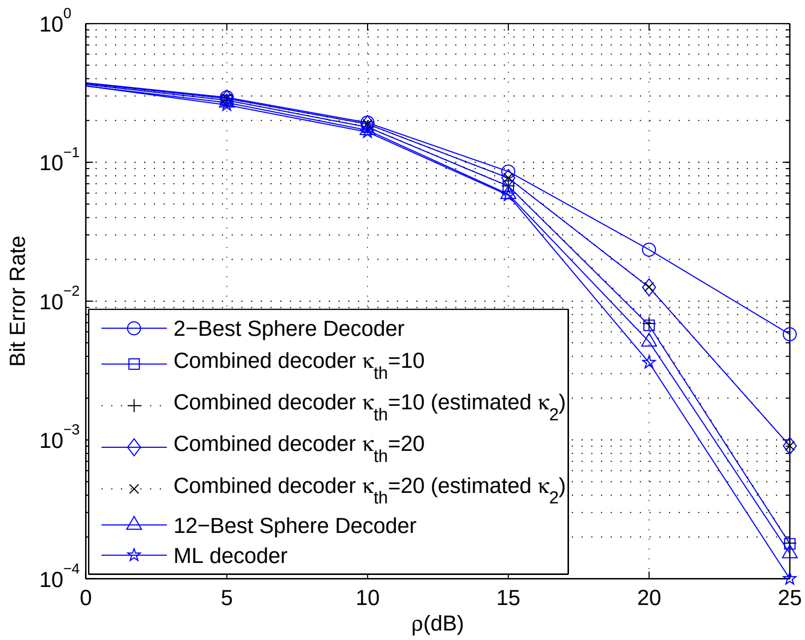

Figure 8.2: Comparison between the BER of the proposed combined K-Best detector with two different thresholds on a 4x4 MIMO using 16QAM, working with the exact and estimated 2-norm condition number, all compared to conventional 2-Best and 12-Best decoders.

Another interesting step was to compare the performance of our combined decoders with the fixed K-Best Sphere Decoders that needed the expansion of the same number of nodes (i.e. the fixed K-Best decoders with the same computational complexity than

Table 8.1: Average number of expanded nodes n for the different decoders under study.

2-Best	3-Best	5-Best	12-Best	Combined $\kappa_{th} = 10$	Combined $\kappa_{th} = 20$
16	24	40	88	37.25	21.89

each of the combined decoders) in order to know if performing a combined detection was worth. In Fig. 8.3, the two cases of combined decoder (with $\kappa_{th} = 10$ and $\kappa_{th} = 20$) are compared to the conventional K-Best SD algorithms that expand the similar average number of nodes, which are respectively 5-Best and 3-Best, as Table 8.1 shows.

It can be observed in Fig. 8.3 that the combined decoders show better performance than the conventional ones, for a given complexity. Note that in Table 8.1 the values of n for the 2-Best and 12-Best SD are also included and, obviously, the values $n_{\kappa_{th}=10}$ and $n_{\kappa_{th}=20}$ remain between them. It can be concluded that exploiting the knowledge about the channel matrix condition number really can help to decrease the detection complexity.

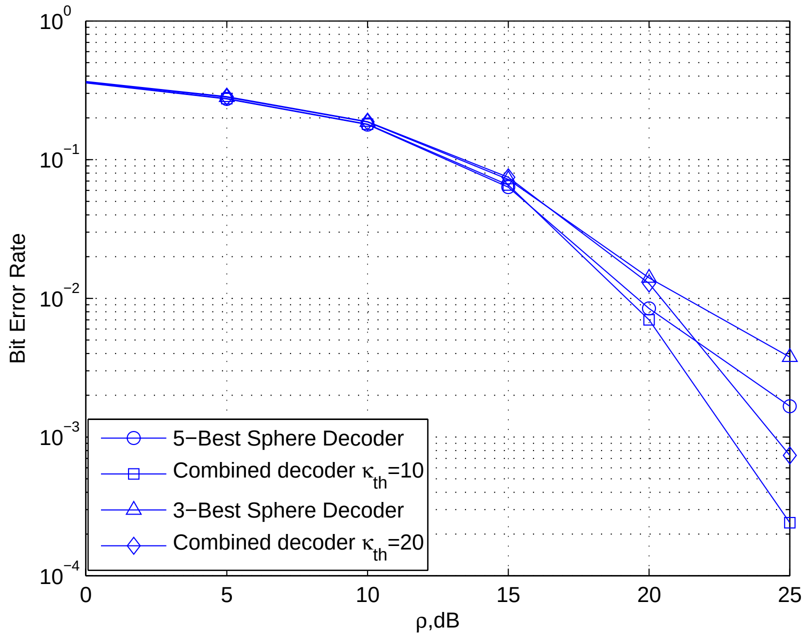

Figure 8.3: BER curves of the proposed combined K-Best detector with two different thresholds on a 4x4 MIMO using 16QAM, both compared to the conventional K-Best decoders with equivalent complexity (3-Best and 5-Best).

8.2 Joint channel estimation and data detection

The performance of the joint ML channel estimation and data detection algorithm is next compared to the one of the disjoint channel estimation and data detection. In the case of the joint channel estimation and detection, the ML detection algorithm used for solving (7.10) is the ASD (detailed in Chapter 2), because by using this method an ML detection can be guaranteed with quite low complexity. Both channel estimation and data detection techniques are also compared to the Perfect CSI case, where the channel is supposed perfectly known at the receiver.

For this purpose, a 1x2 SIMO BPSK system was considered for our simulations. Note that in addition to the data detection, channel estimation is now being carried out, so it will be necessary to send the data in time blocks. For the simulations with the joint channel estimation and detection algorithm, the data is transmitted in blocks of 10 and one training symbol is embedded, resulting in a time period of $T = 11$. For the disjoint channel estimation and detection case, again just one training symbol is used (for consistently comparing with the previous case) and the detection algorithm used is the 1-Best, since this algorithm obtains the ML solution, when there is just one transmit antenna and the symbols belong to a BPSK alphabet, in a very efficient and accurate way. Simulation results are obtained via Monte Carlo runs in which $\mathbf{h}$ and $\mathbf{V}$ are varied, being the entries of both iid complex Gaussian random variables. The comparison among the performance of the three techniques proposed above is depicted in Fig. 8.4.

It can be noted that the joint estimation and detection performs better than the disjoint estimation and detection and the performance achieved by the joint algorithm is quite close to the perfect CSI case.

Another simulation with quasi-ML joint estimation and detection was carried out using the suboptimal algorithms 1-Best and 2-Best. Fig. 8.5 shows that this algorithms can get almost ML results, note that by using a 2-Best detection algorithm the BER curve almost matches the one with the ASD.

Finally, in Fig. 8.6 it can be seen the difference in performance of the joint ML estimation and detection and the perfect CSI case for a SIMO QPSK system with $n_R = 6$ and T=11. It can be noted that the joint channel estimation and data detection algorithm has a worse performance for higher order constellations such as QPSK instead of BPSK.

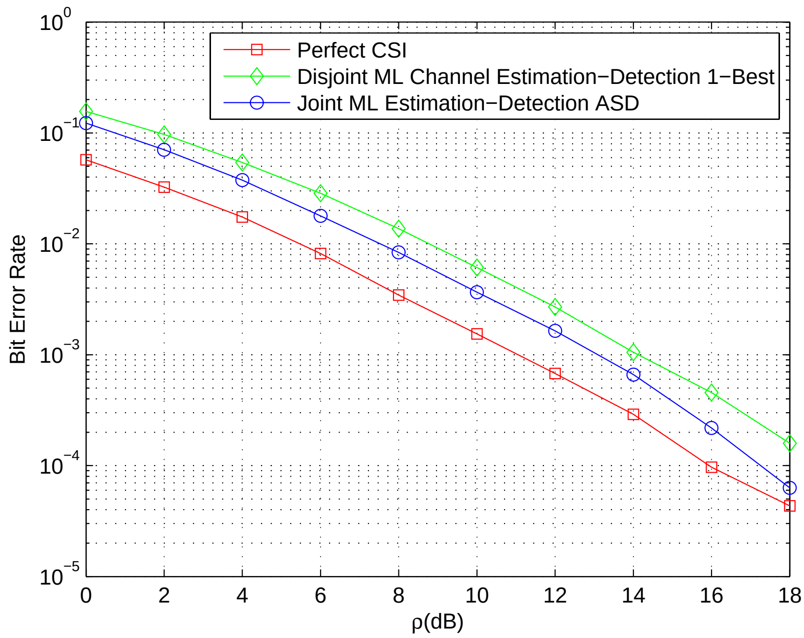

Figure 8.4: BER performance of a SIMO BPSK system with $n_R = 2$ and T=11 in the cases of perfect CSI, disjoint channel estimation and data detection and joint channel estimation and data detection.

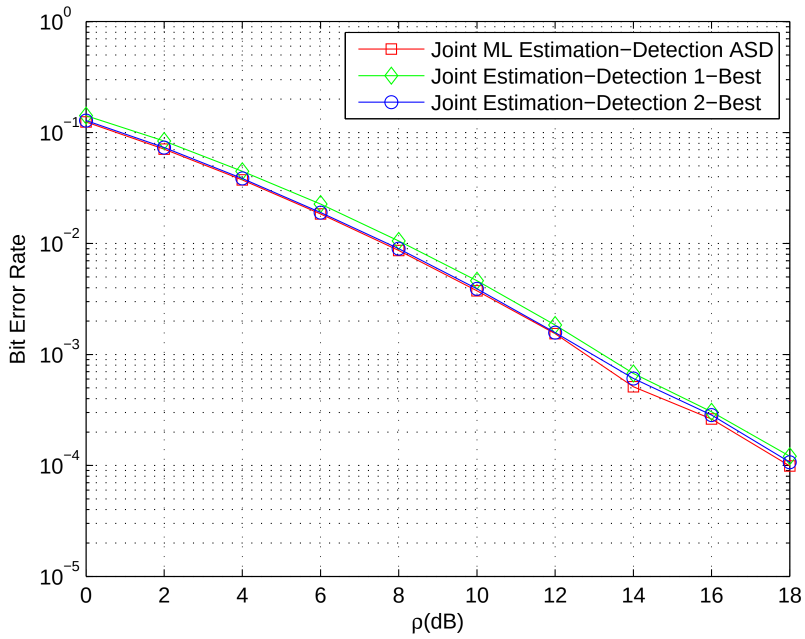

Figure 8.5: BER performance of a SIMO BPSK system with $n_R = 2$ and T=11, with joint channel estimation and data detection using different detectors.

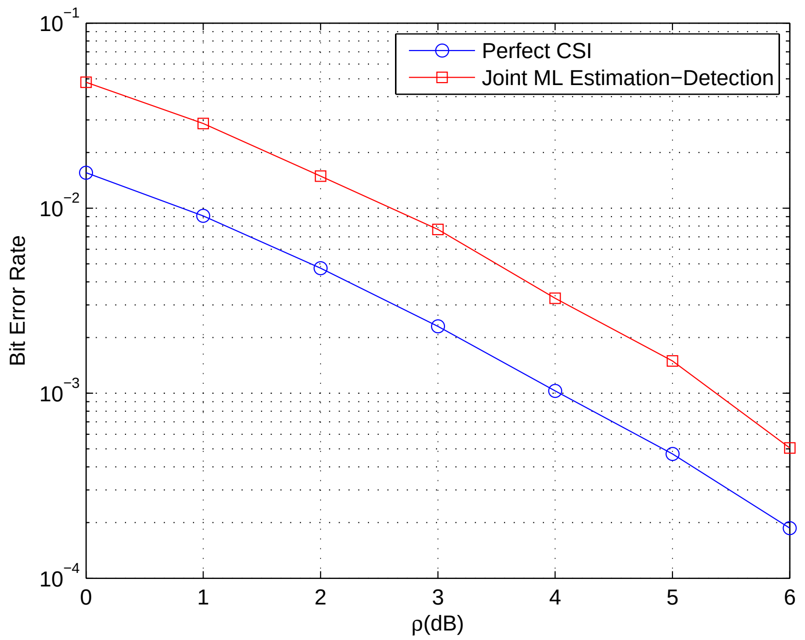

Figure 8.6: BER performance of a SIMO QPSK system with $n_R = 6$ and T=11 in the cases of perfect CSI and joint channel estimation and data detection.

Chapter 9

Summary and Conclusions

Throughout this Master Thesis an overview of detection algorithms for MIMO has been presented. The most important contribution has been the development of a combined decoder that uses different values of K in the K-Best algorithm, depending on the 2-norm condition number of the channel matrix.

It has been shown that before carrying out the combined detection, a channel matrix condition number estimator is needed. For this purpose, a condition number estimator based on the QR decomposition of the channel matrix, which is very useful for combined detectors using SD methods, and also on the Power Method has been proposed. Also, two meaningful ways of determining the threshold condition number to use have also been described. Simulation results have been included in order to discuss the performance of the combined decoder. The BER achieved by two particular cases of combined decoders was represented and it was shown that the proposed decoder together with the condition number estimator and threshold selection algorithms can obtain successful results with the advantage of having bounded complexity.

In the last part, another interesting application of Sphere Decoder methods was presented. It consisted on doing a joint ML channel estimation and data detection for SIMO channels using the optimal Sphere Decoder called ASD. Results show that this method works better than the disjoint channel estimation and data detection even when suboptimal sphere decoders like the 1-Best or 2-Best are employed instead of the ASD. This is an interesting result, since 1-Best and 2-Best have bounded complexity in comparison to ASD and their complexities are known to be much lower. Unfortunately, as the order of the constellation employed increases, the performance of the joint channel estimation and detection method gets worse.

Future work could focus on finding further more efficient ways of carrying out the joint ML channel estimation and data detection and also on achieving better results than the already published for the MIMO case.

Chapter 10

Acknowledgments

First of all, I would like to thank my supervisors Prof. Alberto González and Dr. Vicenç Almenar for his continuous support of my research work at the Audio and Communications Signal Processsing Group (gtac). Without their help, it had been much harder to get through this Master Thesis. My thanks also to Prof. Antonio M. Vidal for his generous advice regarding some computational complexity aspects and for sharing with me his wide mathematical knowledge every time that I need it.

From a personal perspective, I have to mention the support of my parents, thanks for encouraging me so much in every new project of my life. My gratitude to the rest of my family and to all of my friends (classmates, workmates, old and new friends) for every nice moment spent together.

Last, but not least, a big hug to Max for being such a special person.

This work was supported in part by FEDER grant MEC-TEC2006-14204-C02-01.

Bibliography

[1] ARTES, H., SEETHALER, D., AND HLAWATSCH, F. Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection. *IEEE Transactions on Signal Processing 51*, 11 (November 2003), 2808 – 2820.

[2] BARRY, J., LEE, E., AND MESSERSCHMITT, D. *Digital Communications*. Ed. Springer, United States, 2003 (3rd Edition).

[3] BERENGUER, I., AND WANG, X. Space-Time coding and signal processing for MIMO communications. *Journal of Computer Science and Technology 18*, 6 (November 2003), 689–702.

[4] CLINE, A. K., MOLER, C. B., STEWART, G. W., AND WILKINSON, J. H. An estimate for the condition number of a matrix. *SIAM Journal on Numerical Analysis 16*, 2 (April 1979), 368–375.

[5] FINCKE, U., AND POHST, M. Improved methods for calculating vectors for short length in a lattice, including a complexity analysis. *Mathematics of Computation 44* (April 1985), 463–471.

[6] FOSCHINI, G. J. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs. Technical Journal 1*, 2 (Autumn 1996), 41–59.

[7] GENTLE, J. E. *Numerical Linear Algebra for Applications in Statistics*. Ed. Springer, United States, 1998. ISBN: 978-0-387-98542-8.

[8] GOLUB, G., AND LOAN, C. V. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1996.

[9] GUO, Z., AND NILSSON, P. Algorithm and implementation of the K-Best Sphere Decoding for MIMO Detection. *IEEE Journal on Selected Areas in Communications 24*, 3 (March 2006), 491–503.

[10] HASSIBI, B., AND VIKALO, H. On Sphere Decoding algorithm. Part I, the expected complexity. *IEEE Transactions on Signal Processing 54*, 5 (August 2005), 2806–2818.

[11] HOTTINEN, A., KUUSEAL, M., K.HUGL, ZHANG, J., AND RAGHOTHAMAN, B. Industrial embrace of smart antennas and mimo. *IEEE Wireless Communications 13*, 4 (August 2006), 8– 16.

[12] KAILATH, T., VIKALO, H., AND HASSIBI, B. *"MIMO Receive Algorithms," in Space-Time Wireless Systems: From Array Processing to MIMO Communications, (editors H. Bolcskei, D. Gesbert, C. Papadias, and A. J. van der Veen)*. Cambridge University Press, United Kingdom, 2005.

[13] LI, Q., AND WANG, Z. Improved K-Best Sphere Decoding algorithms for MIMO systems. In *International Symposium on Circuits and Systems (ISCAS 2006)* (Island of Kos, Greece, May 2006).

[14] MAURER, J., MATZ, G., AND SEETHALER, D. Low-complexity and full-diversity MIMO detection based on condition number thresholding. In *IEEE International Conference on Audio, Speech and Signal Processing (ICASSP 2007)* (Honolulu, Hawaii, USA, April 2007).

[15] SU, K. Efficient Maximum Likelihood detection for communication over MIMO channels. Technical report, University of Cambridge, February 2005.

[16] VIKALO, H., HASSIBI, B., AND STOICA, P. Efficient joint maximum-likelihood channel estimation and signal detection. *IEEE Transactions on Wireless Communications 5*, 7 (July 2006), 1838 – 1845.

Combined K-Best Sphere Decoder based on the channel matrix condition number

Sandra Roger, Alberto Gonzalez, Vicenc Almenar
Inst. for Telecom. and Multimedia Applications (i-TEAM)
Technical University of Valencia
Valencia, Spain
sanrova@iteam.upv.es, http://www.iteam.upv.es

Antonio Vidal
Dept. of Information Systems and Computation (DSIC)
Technical University of Valencia
Valencia, Spain

Abstract— **It is known that Sphere decoding (SD) methods can provide Maximum-Likelihood (ML) detection over Gaussian MIMO channels with lower complexity than the exhaustive search. Channel matrix condition number represents an important influence on the performance of usual detectors. Throughout this paper, two particular cases of a SD method called K-Best carry out a combined detection in order to reduce the computational complexity with predictable performance degradation. Algorithm selection is based on channel matrix condition number thresholding. K-Best is a suboptimal SD algorithm for finding the ML solution of a detection problem. It is based on a fixed complexity tree search, set by a parameter called *k*. The proposed receiver makes use of a low value of *k* while working with well-conditioned channels and switches to a higher value of *k* whether the channel gets worse. It is also presented practical algorithms for finding the 1-norm condition number of a given channel matrix and the condition number threshold selection. Finally an algorithm variant that switches between an ML SD and a linear detector is also evaluated.**

Keywords- Sphere Decoding, MIMO detection, K-Best; condition number;

I. INTRODUCTION

Motivated by the need of increasingly sophisticated connectivity anytime and anywhere, wireless communications have attracted increasing research attentions recently. Mainly Multiple-Input Multiple-Output (MIMO) wireless communication systems that exhibit several transmit and receive antennas. This feature can provide several advantages, for instance, an increased capacity that can reach, in some scenarios, the Shannon limit. What is really interesting about the benefits offered by MIMO systems is that all of them can be reached without the need for additional spectral resources, which are really expensive and scarce. The most typical ways of using a MIMO system are diversity and multiplexing [1]. This work is focused on the *Bell-Labs Layered Space Time* system (BLAST), an example of multiplexing MIMO system, although the developed algorithm use is not limited to this particular case.

A. Maximum-likelihood decoding

Maximum-likelihood (ML) detection over Gaussian MIMO channels is shown to get the lowest BER for a given scenario [1]. However, it has a prohibitive complexity which grows exponentially with the number of transmit antennas and the size of the constellation, since it makes an exhaustive search to reach the solution. Sphere Decoding (SD) techniques [2,3] can reach the ML solution at lower complexity than the exhaustive search. These methods look for a solution within a hypersphere centered at the received signal vector. Unfortunately, the complexity of regular SD algorithms is strongly dependent on the preprocessing stage to look for the sphere radius and also on its value. SD algorithms computational cost varies also with different signals and channels. Therefore, since the detection throughput is non-fixed, these methods are not suitable for real time detection and hardware implementation. On the other hand, the SD method called K-Best [4] exhibits fixed complexity, but it does not reach the ML solution in all cases.

It has already been shown in [5] that the channel condition number has an important influence on the performance of detectors. Throughout this paper, a K-Best SD decoder, which switches its complexity between a lower or higher value depending on the channel condition number, is developed. Paper presents:

- A combined decoder that uses different values of *k* in the K-Best algorithm depending on the condition number in 1-norm of the channel matrix.
- A condition number estimator based on the QR decomposition of the channel matrix [6].
- A decoder that combines a linear detector and an optimum Sphere Decoder.

II. SYSTEM MODEL

Let us consider a MIMO system with n_T transmit antennas and n_R receive antennas and a signal to noise ratio denoted by ρ. Symbols coming from the data stream are taken in groups of n_T and sent, overlapped in time and frequency, through the n_T transmit antennas. The baseband equivalent model for such MIMO system is given by

$$\mathbf{x} = \sqrt{\frac{\rho}{n_T}}\mathbf{Hs} + \mathbf{v} . \quad (1)$$

Where $\mathbf{s} = \left[s_1, s_2, ..., s_{n_T} \right]^{\mathrm{T}}$ represents the baseband signal vector transmitted during each symbol period, all its elements are chosen from the same constellation such as M-QAM, therefore all possible **s** vectors belong to a finite n_T-dimensional lattice. The constellation points are scaled by a factor so that the average power of the constellation is one. Vector $\mathbf{x} = \left[x_1, x_2, ..., x_{n_R}\right]^T$ in (1) denotes the received symbol vector, and $\mathbf{v} = \left[v_1, v_2, ..., v_{n_R}\right]^T$ stands for an independent identical distributed (i.i.d.) complex zero-mean Gaussian noise vector with unit variance. The channel between each transmit and receive antenna is modeled as flat fading. This way, the channel matrix **H** is formed by $n_R \times n_T$ complex-valued elements, h_{ij}, which represents the fading gain from the *j*-th transmit antenna to the *i*-th receive antenna. The elements of **H** are considered i.i.d. complex zero-mean Gaussian variables with a variance of 0.5 per dimension.

It should be noted that in order to apply Sphere Decoding methods to the system model (1), the complex model is usually transformed into a real one rewriting the complex system as

$$\begin{bmatrix} Re(\mathbf{x}) \\ Im(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} R(\mathbf{H}) & -I(\mathbf{H}) \\ I(\mathbf{H}) & R(\mathbf{H}) \end{bmatrix} \begin{bmatrix} Re(\mathbf{s}) \\ Im(\mathbf{s}) \end{bmatrix} + \begin{bmatrix} Re(\mathbf{v}) \\ Im(\mathbf{v}) \end{bmatrix}. \quad (2)$$

III. SPHERE DECODING ALGORITHMS

Given the received signal, **x**, the detection problem consists in determining the transmitted vector with the highest a posteriori probability. This is typically carried out in practice by means of solving the called least squares problem

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in M^{n_T}} \left\| \mathbf{x} - \sqrt{\frac{\rho}{n_T}} \mathbf{Hs} \right\|^2. \quad (3)$$

Since all the possible **s** vectors belong to a finite n_T-dimensional lattice, a first way of finding the solution of (3) can be simply performing an exhaustive search of M^{n_T} points, which gives a very complex algorithm. Instead of performing an exhaustive search over the total n_T-dimensional lattice points, SD methods [3] limit this search to only the lattice points located within a distance of the received vector lower than a given maximum distance, called sphere radius. The ML solution would then be the closest lattice point of the list of visited points. However, it is necessary to find a suitable value of the sphere radius, what can be difficult in practice.

The search in the SD algorithm is carried out by means of a decoding tree that represents how the distance from the received vector to the solution is calculated as an addition of partial Euclidean distances (PEDs), associated to the tree branches. Partial solutions are represented as nodes and nodes are expanded in order to look for the ML solution, representing a computational effort. It is required to find the ML solution expanding as few nodes as possible. Different tree search strategies have been proposed, some of them can be found in [2,3,4]. A QR factorization of the channel matrix is required to work with a decoding tree structure. This factorization allows to transform the system to an equivalent one that can be solved by means of branch costs and PEDs.

A. K-Best Sphere Decoder

K-Best [4] SD algorithm, instead of expanding every node at each level of the decoding tree or sphere, expands only *k* survivor nodes, which show the smallest accumulated PEDs. It works visiting the tree level by level and expanding only *k* nodes at each level. Finally, when the last level is reached, the leaf node with the smallest total Euclidean distance is selected as solution, see Fig. 1. The detected signal vector is given by the path from the root up to this leaf node.

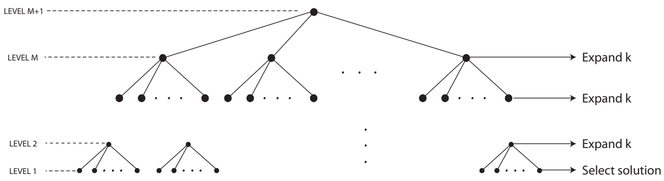

Figure 1. Decoding tree of the K-Best algorithm

The main advantage of this method is that the maximum number of paths is limited and this makes its hardware implementation easier. Variants of this algorithm include a sphere radius in order to reduce the number of explored paths [7] but unfortunately, this number is then non-fixed and unknown. As it is shown in [7], it is more likely to discard the ML solution at early decoding stages, since in latest levels the accumulated PED is closer to the final total distance. Thus, the method can also be modified to work with a different *k* values at different decoding levels, which is called Dynamic K-Best.

IV. CONDITION NUMBER

The sensitivity of the solution of a non-singular system of linear equations $\mathbf{Ax} = \mathbf{b}$ with respect to perturbations of the matrix **A** is directly related to the condition number $\kappa(\mathbf{A})$ [6], which is defined as $\kappa_P(\mathbf{A}) = \|\mathbf{A}\|_P \|\mathbf{A}^{-1}\|_P$. It can be noticed that $\kappa(\mathbf{A})$ is a function of the l_P norm, generally only l_1, l_2 and l_∞ norms are useful in practice. The variation of κ with the norm can be somewhat predicted, since on a finite dimensional vector space, all norms are related [8]. For instance, the κ_1 and κ_2 are related on $\Re^n$ and their equivalence is given by

$$\frac{1}{n}\kappa_2(\mathbf{A}) \le \kappa_1(\mathbf{A}) \le n\kappa_2(\mathbf{A}). \quad (4)$$

A. Condition number estimator

It is important in practice when solving linear systems to have some estimate of $\kappa(\mathbf{A})$ which will give at least a reliable indication of its order of magnitude. When a linear system $\mathbf{Ax} = \mathbf{b}$ has been solved by a direct method, one has some factorization of **A** and it is natural to make use of this in determining the estimate of $\kappa(\mathbf{A})$. The problem is perhaps simpler when we have a factorization of the form **A=QR**, which is known as the QR factorization, where **Q** is orthogonal

$\mathbf{QQ}^T$=**I** and **R** is upper triangular. In this case $\|\mathbf{A}\|_2 = \|\mathbf{R}\|_2$ and $\|\mathbf{A}^{-1}\|_2 = \|\mathbf{R}^{-1}\|_2$, what means that $\kappa_2(\mathbf{A}) = \kappa_2(\mathbf{R})$. However, we concentrate on estimating $\kappa_1(\mathbf{R})$, since the l_1 norm of **R** can be efficiently computed [6][9] and can be used as an estimator of $\kappa_1(\mathbf{A})$. Although other possibilities exist, this estimator has shown to be very useful when working with SD methods, since a QR factorization of the channel matrix is always available and it is used in the present work.

V. Combined Decoder

As it was explained above, the proposed K-best decoder will select a low value of *k* while working with well-conditioned channels and switch to a higher value of *k* whether the channel is unfortunately ill-conditioned. This way a greater decoding tree is visited when dealing with poor channels and there is less probability of discarding the ML solution too early. An estimator of the condition number together with a threshold condition number, denoted by κ_{th}, are used to classify the channels and consequently fix the *k* parameter.

In order to estimate κ_{th} it is very helpful to represent within a determined range of condition numbers the BER of the lower performance algorithm used in the combined decoder. This way it can be observed the condition number values that make the BER exceed the mean BER achieved by this algorithm. For instance, the mean BER of the K-Best algorithm with *k*=2 and $\rho = 20dB$ on a 4x4 MIMO system working with a 16-QAM alphabet was given by $\text{BER}_{2\text{-Best}}$=0.022. As expected, Fig. 2 shows that the BER of this decoder increases as the condition number of the matrix does. The BER is almost every time below the $\text{BER}_{2\text{-Best}}$, for channels with κ<30, see Fig. 2, therefore this value of κ_{th} could be chosen. In section IV, the value of $\kappa_{th} = 60$ is selected as well for this system, in order to compare the results.

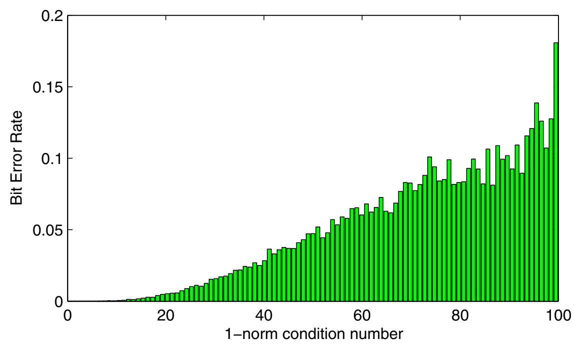

Figure 2. BER for the K-Best decoder with *k*=2 and $\rho = 20dB$ on a 4x4 MIMO and 16 QAM as a function of the channel condition number in 1-norm.

An alternative to choose the threshold condition number can be based on the desired average computational cost of the combined algorithm, which is related to the average number of expanded nodes, *n*. It is somewhat possible to estimate the threshold condition number using the desired average number of expanded nodes, *n*, and the number of expanded nodes of each of the decoders that compose the combined one, together with the probability distribution of the 1-norm condition number of a given random model of channel matrix [10]. For instance, in case of combining two K-Best algorithms with k_1 and k_2, $k_1 < k_2$, for a threshold value of κ_{th}, the relationship between the numbers of expanded nodes, which is very closely related to the computational costs, would be

$$n_{\kappa=\kappa_{th}} = (1 - P_{\kappa\geq\kappa_{th}}) \cdot n_{k_1-\text{Best}} + P_{\kappa\geq\kappa_{th}} \cdot n_{k_2-\text{Best}}. \quad (5)$$

where $P_{\kappa\geq\kappa_{th}}$ is the probability of having a channel with condition number higher than a threshold κ_{th}.

VI. Results

In this section, we discuss the performance of the proposed combined receiver by means of simulations. The particular case of MIMO system used for simulations was a 4x4 system, working with a 16-QAM alphabet. In all the simulations the number of different realizations of the Gaussian channel was between 100 and 1000. The combined receiver switches between two K-Best algorithms: one with *k*=2, where the decoding tree is pruned from the first level, and the other with *k*=12, where the tree is pruned from the second level. Obviously, the behavior of the last case is almost ML. The performance of the combined receiver with condition number thresholds of $\kappa_{th} = 30$ and $\kappa_{th} = 60$ are shown in Fig. 3 and compared to an ML decoder. As expected, the performance is poorer as the threshold increases. Fig. 3 also includes the effect of the use of the condition number estimator instead of the exact value for the two cases of the combined receiver. It can be noticed that the estimation of the condition number does not degrade the performance of the detector in both cases.

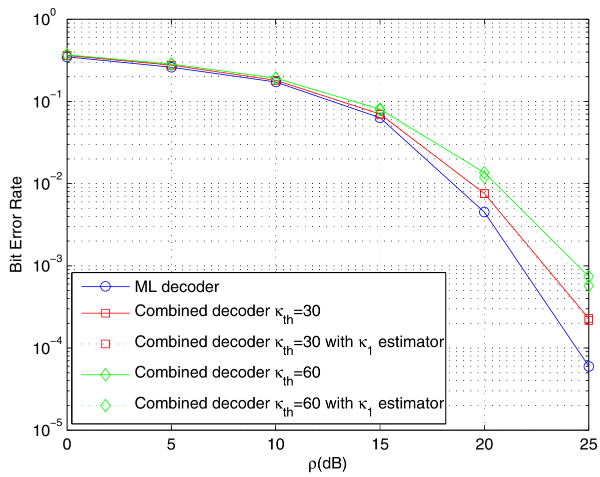

Figure 3. BER curve of the proposed combined K-best detector with two different thresholds on a 4x4 MIMO and 16 QAM, using the exact and estimated 1-norm condition number, all compared to ML detection.

Table I shows the average number of expanded nodes calculated using (5) for a K-Best decoder with k=2 and k=12, and for combined decoders with $\kappa_{th} = 30$ and $\kappa_{th} = 60$. It can be noticed that for a given threshold, the complexity of a combined decoder can be predicted. And the other way round, a desired complexity can set the most suitable threshold. On the other hand, the complexity of the combined decoder reduces to a half for $\kappa_{th} = 30$ or to a quarter for $\kappa_{th} = 60$ compared to the complexity of the 12-Best. The SNR degradation for BER=10^{-3} lies between 1 and 3dB, see Fig. 3.

TABLE I. AVERAGE EXPANDED NODES

Average number of expanded nodes n			
2-Best	12-Best	Combined $\kappa_{th} = 30$	Combined $\kappa_{th} = 60$
16	88	40.06	23.41

Finally, a variant of this combined receiver is presented. In this case the chosen low performance detector is ZF-SIC [1], a linear detector, and the K-Best detector with the higher value of k is replaced by an SD method called Automatic Sphere Decoder (ASD) [11] which always reaches the ML solution. The thresholds $\kappa_{th} = 10$ and $\kappa_{th} = 30$ were again chosen as described in section V. Notice that these thresholds are lower than the ones in the combined decoder using K-Best algorithms. This was somewhat expected since the condition number was calculated over the complex channel matrix (4x4) with half of the dimension of the real one and it is known that when the matrix elements are randomly and independently distributed as normal or uniform with zero mean and unit variance, the increase of the condition number is approximately linear in the size of the matrix [10]. Fig. 4 illustrates the performance of this detector. As it happened in Fig. 3, the higher the threshold is, the poorer the detector performs. The advantage of this method is that now we have a method that is ML and therefore achieves the best performance. Unfortunately, now neither the complexity can be bounded by a value nor the hardware adapted to deal with a single type of an scalable algorithm, as it occurs when using K-Best algorithms.

VII. CONCLUSIONS

A combined receiver that switches between a better performing decoder and worse, but faster, one, has been developed. The criterium to switch is based on the condition number of the channel matrix. In particular, two different combined decoders were presented, one switching between two K-Best algorithms and another combining ASD and ZF-SIC decoders. A method to compute the 1-norm condition number of a matrix using the QR decomposition, which has to be necessarily calculated for Sphere Decoding, is also presented together with two ways for selecting the condition number threshold, which have shown to be very useful in practice. In the simulations, we could observe that the final complexity can be adjusted with a threshold value in the K-Best combined decoder. Another interesting result is that the condition number estimation does not alter the achieved BER of a combined detector. Both things make the K-best based combined decoder a meaningful detection algorithm for practical systems.

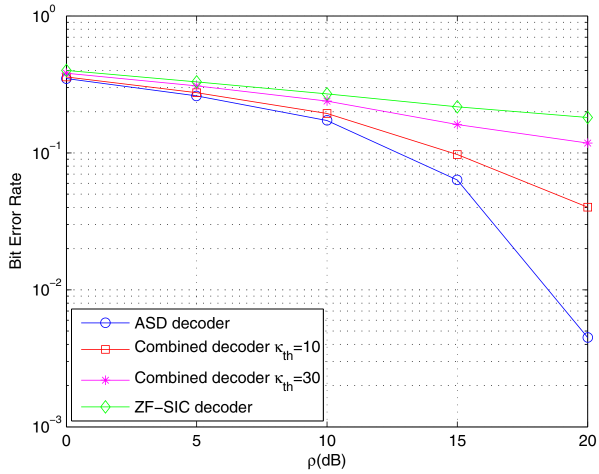

Figure 4. BER curve of the ASD, ZF-SIC decoders and a combination of them with different thresholds on a 4x4 MIMO and 16 QAM.

ACKNOWLEDGMENT

Thanks to Jose Marin-Roig for all his help to improve our K-Best algorithm software. This work was supported by FEDER-MEC under grant TEC2006-14204-C02-01.

REFERENCES

[1] I. Berenguer and X. Wang, "Space-time coding and signal processing for MIMO communications," Journal of Computer Science and Technology, vol. 18, Issue 6, pp. 689-702, November 2003.

[2] B. Hassibi and H. Vikalo, "On sphere decoding algorithm. Part I, The expected complexity," IEEE Transactions on Signal Processing, vol. 54, Issue 5, pp. 2806-2818, August 2005.

[3] U. Fincke and M. Pohst, "Improved methods for calculating vectors for short length in a lattice, including a complexity analysis," Mathematics of Computation, vol. 44, pp. 463-471, April 1985.

[4] Z. Guo and P. Nilsson, "Algorithm and Implementation of the K-Best Sphere Decoding for MIMO Detection," IEEE Journal on Selected Areas in Communications, vol. 24, no. 3, pp. 491-503, March 2006.

[5] J. Maurer, G. Matz and D. Seethaler. "Low-complexity and full-diversity MIMO detection based on condition number thresholding," Proceedings of ICASSP 2007, vol.3, pp. III-61-64. Honolulu, Hawaii, April 2007.

[6] A. K. Cline, C. B. Moler, G. W. Stewart, J. H. Wilkinson, "An Estimate for the Condition Number of a Matrix," SIAM Journal on Numerical Analysis, vol. 16, no. 2 , pp. 368-375, April 1979.

[7] Q. Li and Z. Wang, " Improved K-Best Sphere Decoding algorithms for MIMO systems," International Symposium on Circuits and Systems (ISCAS 2006), Island of Kos, Greece. 21-24 May 2006.

[8] S. Xu and J. Zhang, "A new data mining approach to predicting matrix condition numbers," Communications in information and systems, International Press, vol. 4, no. 4, pp. 325-340, 2004.

[9] J.J. Dongarra, J.R. Bunch, C.B. Moler and G.W. Stewart, "LINPACK Users' Guide", SIAM, Philadelphia, 1979.

[10] J. E. Gentle, "Numerical Linear Algebra for Applications in Statistics," Ed. Springer 1998.

[11] K. Su, "Efficient Maximum Likelihood detection for communication over MIMO channels," Technical report, University of Cambridge, February 2005.

MIMO Channel Matrix Condition Number Estimation and Threshold Selection for Combined K-Best Sphere Decoders

Sandra Roger, Alberto Gonzalez, Vicenc Almenar and Antonio M. Vidal

***Abstract*—It is known that channel matrix condition number represents an important influence on the usual detectors performance. Therefore, a different detection algorithm depending on the channel matrix condition number can be selected in order to achieve a lower complexity than already proposed algorithms with similar performance. Several authors have proposed combined decoders based on channel matrix condition number thresholding. These combined algorithms need an estimation stage of the channel matrix condition number and a previous selection of a suitable threshold condition number. This letter presents a meaningful algorithm for finding the 2-norm condition number of a MIMO channel matrix, specially suitable for combined sphere decoders. Also, two possible threshold selection methods are presented. A practical implementation of a combined K-Best decoder is shown as application example.**

***Index Terms*—MIMO detection, K-Best Sphere Decoder, matrix condition number estimation, threshold selection.**

I. Introduction

Maximum-likelihood (ML) detection over Gaussian Multiple-Input Multiple-Output (MIMO) channels is shown to get the lowest Bit Error Rate (BER) for a given scenario [1]. However, it has a prohibitive complexity which grows exponentially with the number of transmit antennas and the size of the constellation. Motivated by this, there is a continuous search for computationally efficient suboptimal detectors, as the well-known linear detectors based on the ZF or MMSE approaches [1]. Recently, some other suboptimal techniques as the K-Best Sphere Decoder (SD) algorithm have been developed [2], [3]. These methods exhibit fixed complexity, which is very useful for real time detection and hardware implementation [4]. Furthermore, experiments [5] show that the channel matrix condition number is strongly related to the performance of these suboptimal detection schemes, since it is a measure of how the original constellation is distorted by the channel. For instance, Fig.1 shows the degradation performance of some MIMO detectors with the channel matrix condition number for a value of $\rho = 20$. Note that it can be found a channel matrix condition number value that makes the performance of 5-Best equal the one of 8-Best or even the one of ML. Also, as the condition number gets higher, the performance of 8-Best is not as much degraded as the one of 5-Best. Considering this, it seems obvious that a suitable combination of both algorithms based on the channel condition number could have an almost ML behavior for this value of ρ. It can also be shown [6] that the detection degradation increases with the number of antennas, which fixes the size of the channel matrix.

Sandra Roger (1), Alberto Gonzalez and Vicenc Almenar are with the Institute of Communications and Multimedia Technologies (iTEAM), Technical University of Valencia, Spain e-mail: (1) sanrova@iteam.upv.es.

Antonio M. Vidal is with the Department of Computer Systems and Computation (DSIC), Technical University of Valencia, Spain.

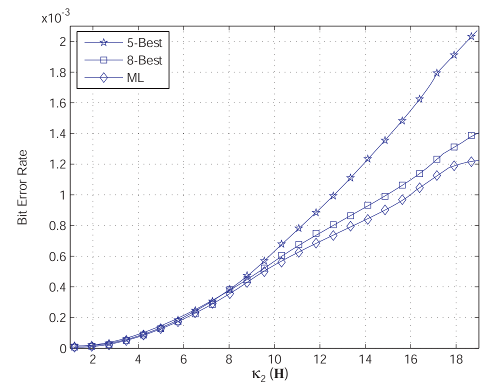

Fig. 1. BER of different MIMO detectors with a 16-QAM constellation, a 4x4 MIMO channel and a value of $\rho = 20$, as a function of the channel matrix condition number $\kappa_2(\mathbf{H})$.

Several authors have developed combined detectors based on condition number thresholding, for instance ML and ZF were combined in [7] and two K-Best algorithms with different K were combined in [8]. These combined detectors need an estimation stage of the channel matrix condition number and a previous selection of the most suitable threshold condition number in each case, as illustrated by Fig.1. This letter presents first a meaningful algorithm for finding the 2-norm condition number and next two possible threshold selection strategies. Although the proposed methods are specially suitable for combined Sphere Decoders, they can also be useful for other combined detectors.

A. System model.

Present work is focused on the well-known Bell-Labs Layered Space Time system (BLAST), although its contribution is not limited to this particular case. Let us consider a MIMO system with n_T transmit antennas, n_R receive antennas with $n_R \geq n_T$ and a signal to noise ratio denoted by ρ. The baseband equivalent model for such MIMO system is given by

$$\mathbf{x_c} = \mathbf{H_c}\mathbf{s_c} + \mathbf{v_c}, \qquad (1)$$

where $\mathbf{s_c}$ represents the baseband signal vector transmitted during each symbol period formed by elements chosen from the same constellation such as M-QAM. Vector $\mathbf{x_c}$ in (1) denotes the received symbol vector and $\mathbf{v_c}$ is a complex white Gaussian noise vector with zero mean and unit variance. The channel matrix $\mathbf{H_c}$ is modelled as flat fading and it is formed by $n_R \times n_T$ complex-valued elements, h_{ij}, which represent the complex fading gain from the j-th transmit antenna to the i-th receive antenna. Moreover, the channel matrix $\mathbf{H_c}$ is considered known at the receiver and block fading. It should be noted that in order to apply SD methods to the system model (1), the complex model is usually transformed into a real one [2] and this fact will affect to the condition number value, since it depends on the channel matrix size. From now on, the real form of the system (1) will be considered, where the real signal vectors will become $\mathbf{x}$, $\mathbf{s}$, $\mathbf{v}$ and the real channel matrix will be now called $\mathbf{H}$.

B. K-Best Sphere Decoding Algorithms.

Given the received signal $\mathbf{x}$, the detection problem consists in determining the transmitted vector $\hat{\mathbf{s}}$ with the highest a posteriori probability. This is typically carried out in practice by means of solving the following least squares problem

$$\hat{\mathbf{s}} = \arg \min_{s \epsilon M^{2n_T}} ||\mathbf{x} - \mathbf{Hs}||^2 , \qquad (2)$$

where $\|\cdot\|$ denotes the 2-norm. A QR factorization of the channel matrix ($\mathbf{H} = \mathbf{QR}$) allows transforming the system to an equivalent one that can be solved using a tree structure [9]. Matrix $\mathbf{Q}$ is orthogonal, $\mathbf{QQ^T} = \mathbf{I}$, and matrix $\mathbf{R}$ can be decomposed in an upper triangular $2n_T \times 2n_T$ matrix, denoted by $\mathbf{R}_1$, and a $(2n_R - 2n_T) \times 2n_T$ matrix of zeroes. In case of multiplying (2) by $\mathbf{Q^T}$ and calling $\mathbf{y} = \mathbf{Q^T x}$, the problem (2) can be equivalently expressed as

$$\begin{aligned} \hat{\mathbf{s}} &= \arg \min_{s \epsilon M^{2n_T}} ||\mathbf{y} - \mathbf{R_1 s}||^2 & (3) \\ &= \arg \min_{s \epsilon M^{2n_T}} \sum_{i=1}^{2n_T} |y_i - \sum_{j=i}^{2n_T} r_{ij} s_j|^2. & (4) \end{aligned}$$

where the triangular structure of $\mathbf{R_1}$ has also been exploited.

In order to solve (4) via a tree search, the following recursion is performed for $i = 2n_T, \dots, 1$:

$$T_i(S^{(i)}) = T_{i+1}(S^{(i+1)}) + |e_i(S^{(i)})|^2 \qquad (5)$$

$$e_i(S^{(i)}) = y_i - \sum_{j=i}^{2n_T} r_{ij} s_j , \qquad (6)$$

where i denotes each tree level, $S^{(i)} = [s_i, s_{i+1}, \dots, s_{2n_T}]$, $T_i(S^{(i)})$ is the accumulated Partial Euclidean Distance (PED) up to level i and $|e_i(S^{(i)})|^2$ is the distance between levels i and $i+1$ in the decoding tree, which will be represented as the weight of branch. Partial solutions are represented as nodes n and nodes are expanded in order to look for the ML solution or the closest lattice point. It is required to find the ML solution expanding as few nodes as possible in order to reduce the computational effort. K-Best SD algorithm [2] expands only those K survivor nodes that show the smallest accumulated PEDs at each level of the decoding tree. The detected signal vector $\hat{\mathbf{s}}$ is given by the path from the root up to the leaf node with the smallest total Euclidean distance. The main advantage of this method is that the maximum number of visited paths is limited, that yields a fixed computational effort and makes the algorithm hardware implementation easier.

C. Combined K-Best Sphere Decoder.

In order to better illustrate the contributions of this letter, it will be considered the Combined K-Best Sphere Decoder proposed in [8]. Fig.2 depicts the flow diagram of this combined sphere decoder which always works with a K-Best detector but it adjusts its K value depending on how the channel is conditioned, a low value of K is used for well-conditioned channels and a higher value of K is selected whether the channel is unfortunately ill-conditioned. This way, a more complex detection algorithm is used when dealing with poor channels, thus there is less probability of discarding the ML solution too early. The steps to perform the combined detection are next detailed. Firstly a threshold condition number is chosen. It must be checked whether $\mathbf{H}$ changed, in case it did not change, the currently fixed K-Best is used, otherwise the new channel condition number has to be estimated. Then, using the threshold κ_{th}, the suitable K value is selected between k_1 and k_2 with $k_2 > k_1$. Finally, K-Best Sphere Decoding is used for carrying out the detection. It can be observed in Fig.2 that, as said before, an estimator of the condition number together with a threshold condition number, denoted by κ_{th}, have been used to classify the channels and consequently fix the K value. This letter develops a low complexity estimator of the condition number of a matrix that makes use of the QR factorization [10], which is always available when working with SD methods [9] and often with other suboptimal detectors [1]. A meaningful estimator of the 2-norm of a matrix together with the Power Method for computing eigenvalues [10] can provide a reliable and useful approximation of the condition number of the channel matrix.

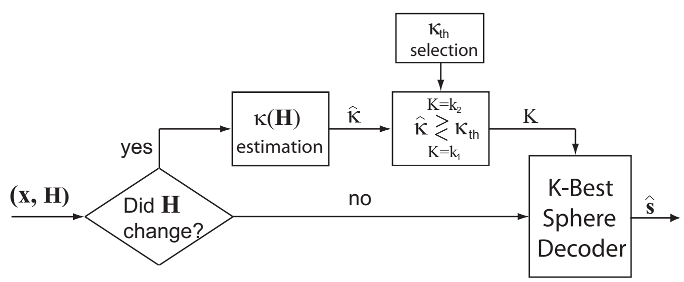

Fig. 2. Flow diagram of a combined K-Best Sphere Decoder with threshold selection and channel condition number estimation.

II. Condition number estimation.

The sensitivity of the solution of a non-singular system of linear equations $\mathbf{Ax} = \mathbf{b}$ with respect to perturbations of the matrix $\mathbf{A}$ is directly related to its condition number [11]. Although the matrix condition number depends on the selected norm, if the matrix is well-conditioned, the condition number will be small in all norms, otherwise it will be large. Thus, the most convenient norm is usually selected between the 1-norm, 2-norm and ∞-norm. The 2-norm condition number is defined as

$$\kappa_2(\mathbf{A}) = \frac{\sigma_{max}}{\sigma_{min}}, \tag{7}$$

being σ_{max} and σ_{min} the maximum and minimum singular values of $\mathbf{A}$ respectively. When $\mathbf{A}$ is square $n \times n$, $\kappa_2(\mathbf{A})$ can be also computed as $\|\mathbf{A}\|\|\mathbf{A^{-1}}\|$. Although other condition numbers can be considered, $\kappa_2(\mathbf{A})$ will be selected in our work because of some special properties presented below. Fig.3 shows the probability density function of the 2-norm condition number for a 8x8 real Gaussian MIMO channel matrix, it can be seen that although lower condition number values are more likely, there also appear several high values of it that can degrade the performance of suboptimal detectors.

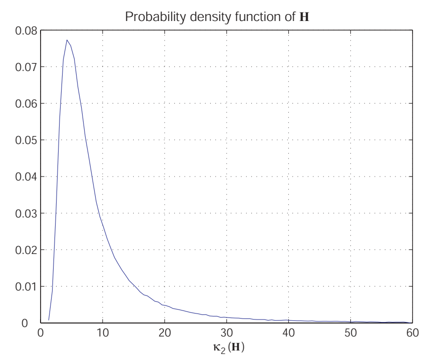

Fig. 3. Experimental probability density function of the 2-norm condition number of 8x8 real Gaussian channel matrices.

In order to carry out the combined detection in practice, it is important to have a reliable estimate of $\kappa_2(\mathbf{H})$. As explained above, K-Best requires a factorization of the form $\mathbf{H} = \mathbf{QR}$. Only in case of working with 2-norm it is true that

$$\kappa_2(\mathbf{H}) = \kappa_2(\mathbf{R_1}) = \|\mathbf{R_1}\|\|\mathbf{R_1^{-1}}\|. \tag{8}$$

Due to the fact that $\mathbf{R_1}$ is triangular, it can be noted that $\|\mathbf{R_1}\|$ can be calculated faster than $\|\mathbf{H}\|$. Moreover, $\|\mathbf{R_1}\| = \sigma_{max}$ can be efficiently computed by applying the Power Method [10]. Furthermore, it will be possible to avoid calculating the inverse of $\mathbf{R_1}$, which requires $O(n^3)$ operations, by using an appropriate estimator of $\|\mathbf{R_1^{-1}}\| = 1/\sigma_{min}$ with only $O(n^2)$.

A. *The Power Method for computing* $\|\mathbf{R_1}\|$

The Power Method is an iterative algorithm that obtains the largest eigenvalue of a given matrix. Given a $n \times n$ diagonalizable matrix $\mathbf{A}$ with $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$ eigenvalues, this method starts with a unit 2-norm vector $\mathbf{q}^{(0)} \in \mathcal{R}^n$ as an initial approximation of one of the dominant eigenvectors. At each iteration i, it computes the new $\mathbf{q}^{(i)}$ in two steps. First, the vector $\mathbf{z}^{(i)} = \mathbf{A}\mathbf{q}^{(i-1)}$ is calculated and next it is normalized resulting in $\mathbf{q}^{(i)} = \mathbf{z}^{(i)}/\|\mathbf{z}^{(i)}\|$. After the last iteration of the process, the maximum eigenvalue can be computed as

$$\lambda_{max} = [\mathbf{q}^{(i)}]^T \mathbf{A}\mathbf{q}^{(i)}. \tag{9}$$

In our work, the Power Method is proposed for calculating $\|\mathbf{R_1}\| = \sigma_{max}$, previously calculating the maximum eigenvalue of $\mathbf{A} = \mathbf{R_1^T R_1}$ which corresponds to σ^2_{max} (obviously, the method can be applied to $\mathbf{R_1^T R_1}$ without computing explicitly this product). Considering that the maximum size of the channel matrices is usually up to 8x8, a maximum number of 10 iterations for running the Power Method gets quite accurate results. Thus, the number of flops can be computed as $21n^2 + 22n$.

B. *Estimator of* $\|\mathbf{R_1^{-1}}\|$

This method was firstly developed in [11]. Two triangular systems need to be formulated. The first one is $\mathbf{R_1^T x} = \mathbf{b}$ and $\mathbf{b}$ has to be chosen so that its solution $\hat{\mathbf{x}}$ will make $\|\hat{\mathbf{x}}\|/\|\mathbf{b}\|$ as large as possible. This is achieved by an iterative process with n steps, considering the size of $\mathbf{R_1^{-1}}$ is n. At each step i, b_i is chosen between $+1$ and -1, in order to maximize x_i, which will be computed as

$$r_{ii}x_i = b_i - (r_{1i}x_i + \ldots + r_{i-1i}x_{i-1}). \tag{10}$$

The second system to solve is $\mathbf{R}_1\mathbf{y} = \hat{\mathbf{x}}$. Once its solution $\hat{\mathbf{y}}$ is obtained, the estimation for $\|\mathbf{R_1^{-1}}\| = 1/\sigma_{min}$ is given by $\|\hat{\mathbf{y}}\|/\|\hat{\mathbf{x}}\|$. Although other estimators of $1/\sigma_{min}$ have been proposed, for example in [7] the authors estimate both σ_{max} and $1/\sigma_{min}$ by means of the Power Method, this estimator has been selected in the present work mainly because it exploits the QR factorization used in the K-Best algorithms. Also it is very suitable to be implemented in practice and it only requires $2n^2 + 6n$ flops.

As soon as σ_{max} and $1/\sigma_{min}$ are available, $\kappa_2(\mathbf{R_1})$ is calculated as the product of both values. Fig.**??** shows that the relative error of our proposed estimator for $\kappa_2(\mathbf{R_1})$ is higher than the one of the Power Method for computing the whole condition number. However, as can be seen in Fig.**??**, the error magnitude is not very significant, making this estimator useful for combined detectors. Table 1 shows that the complexity of our proposed estimator measured in number of flops is almost the half of the complexity of the estimator that only employs the Power Method.

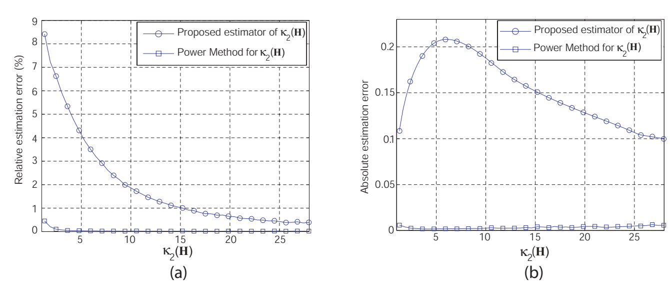

Fig. 4. Error of our proposed estimator compared with the error of the Power Method for computing the whole condition number. (a) Relative. (b) Absolute.

III. Threshold selection

Two meaningful ways of selecting the threshold condition number κ_{th} in a combined decoder are presented throughout

TABLE I
MEASURED COMPLEXITY IN NUMBER OF FLOPS

Real channel	Proposed estimator for $\kappa_2(\mathbf{R_1})$	Power Method for $\kappa_2(\mathbf{R_1})$
4x4	482	850
8x8	1698	3042
16x16	6338	11458

this section. The first one is based on the analysis of the achieved BER of the lower performance algorithm used in the combined decoder. Considering the algorithm has an average BER denoted by BER_{av}, a suitable κ_{th} must guarantee that $\mathrm{BER} \leq \mathrm{BER}_{av}$ when our channel has $\kappa_2(\mathbf{H}) \leq \kappa_{th}$. This will be done in order to assure that the lower performance algorithm is only used when it has optimal performance. Taking into account these considerations, BER values of the lower performance algorithm are represented within a range of condition numbers in a similar manner than in Fig.1. The condition number value corresponding to BER_{av} is selected as κ_{th}, since for $\kappa_2(\mathbf{H})$ values higher than κ_{th} it becomes true that $\mathrm{BER} > \mathrm{BER}_{av}$.

The second method for selecting the threshold states that the chosen κ_{th} should provide the desired average computational cost of the combined algorithm, which is related to the average number of expanded nodes *n*. For instance, in case of combining k_1-Best SD and k_2-Best SD, considering $k_1 < k_2$, for a threshold value of κ_{th}, the resulting average number of expanded nodes $n_{\kappa_{th}}$ would be given by

$$n_{\kappa_{th}} = (1 - P_{\kappa \geq \kappa_{th}}) n_{k_1-Best} + P_{\kappa \geq \kappa_{th}} n_{k_2-Best}, \quad (11)$$

where $P_{\kappa \geq \kappa_{th}}$ is the probability of having a channel with condition number higher than a threshold κ_{th} and it can be calculated as a cumulative distribution of the probability density function depicted in Fig.3. The number of expanded nodes in each k_i-Best algorithm is denoted by n_{k_i-Best} (for $i = 1, 2$). It can be noted that the desired maximum average number of expanded nodes will determine the threshold value. In the same way, for a given threshold, the number of expanded nodes will be straightforwardly predicted.

IV. RESULTS

A 4x4 MIMO system and a 16-QAM alphabet were considered for our simulations. Our proposed condition number estimator and threshold selection methods were applied over a combined K-Best algorithm with $k_1 = 2$ and $k_2 = 12$. The threshold values $\kappa_{th} = 10$ and $\kappa_{th} = 20$ were chosen following the first proposed threshold selection method. The algorithm was run first using the exact 2-norm condition number and afterwards using its proposed estimator. Fig.5 illustrates that estimating the 2-norm condition number does not change the performance of the combined detector significantly. Therefore, the estimator is very suitable for this combined sphere decoder.

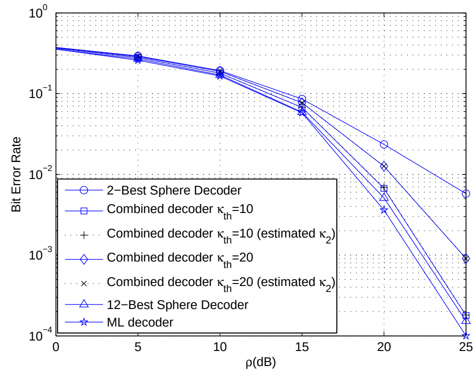

Fig. 5. Comparison between the BER of the combined K-Best detector with two different thresholds on a 4x4 MIMO using 16QAM, working with the exact and estimated 2-norm condition number.

V. CONCLUSION

Throughout this paper it has been proposed a condition number estimator based on the QR decomposition of the channel matrix and also on the Power Method, which has been shown to be very useful for combined detectors using SD methods. Moreover, two meaningful ways of determining a suitable threshold were also cited. In the last part, two particular cases of combined detectors were simulated, using in both the exact and estimated condition number. It can be concluded that the condition number estimator does not degrade the performance of the combined sphere decoder at all and the threshold selection algorithms provide successful results.

REFERENCES

[1] I. Berenguer and X. Wang, "Space-Time coding and signal processing for MIMO communications," *Journal of Computer Science and Technology*, vol. 18, no. 6, pp. 689–702, November 2003.

[2] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-Best Sphere Decoding for MIMO Detection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 491–503, March 2006.

[3] M. Shabany, K. Su, and P. G. Gulak, "A pipelined scalable high-throughput implementation of a Near-ML K-Best complex lattice decoder," in *IEEE International Conference on Audio, Speech and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, USA, April 2008.

[4] M. Wenk, M. Zellweger, A. Burg, N. Felber, and W. Fichtner, "K-best MIMO detection VLSI architectures achieving up to 424 Mbps," in *IEEE International Symposium on Circuits and Systems (ISCAS 2006)*, Island of Kos, Greece, May 2006.

[5] H. Artes, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: A geometrical approach to approximate ML detection," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2808 – 2820, November 2003.

[6] J. E. Gentle, *Numerical Linear Algebra for Applications in Statistics*. United States: Ed. Springer, 1998, iSBN: 978-0-387-98542-8.

[7] J. Maurer, G. Matz, and D. Seethaler, "Low-complexity and full-diversity MIMO detection based on condition number thresholding," in *IEEE International Conference on Audio, Speech and Signal Processing (ICASSP 2007)*, Honolulu, Hawaii, USA, April 2007.

[8] S. Roger, A. Gonzalez, V. Almenar, and A. M. Vidal, "Combined K-Best Sphere Decoder based on the channel matrix condition number," in *International Symposium on Communications, Control and Signal Processing (ISCCSP 2008)*, St. Julians, Malta, March 2008.

[9] B. Hassibi and H. Vikalo, "On Sphere Decoding algorithm. Part I, the expected complexity," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 2806–2818, August 2005.

[10] G. Golub and C. V. Loan, *Matrix Computations*. Baltimore: The Johns Hopkins University Press, 1996.

[11] A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, "An estimate for the condition number of a matrix," *SIAM Journal on Numerical Analysis*, vol. 16, no. 2, pp. 368–375, April 1979.