Lattice Reduction-Aided Minimum Mean Square Error *K*-Best Detection for MIMO Systems

Sébastien Aubert[1,3], Youssef Nasser[2], and Fabienne Nouvel[3]

[1]Eurecom, 2229, route des Crêtes, 06560 Sophia-Antipolis, France

[2]American University of Beirut, ECE Department, Bliss Street, Beirut, Lebanon

[3]Université Européenne de Bretagne, France / INSA-IETR, UMR 6164, F-35708 Rennes, France

Email: sebastien.aubert@eurecom.fr, youssef.nasser@aub.edu.lb, fabienne.nouvel@insa-rennes.fr

ABSTRACT- *The Multiple-Input Multiple-Output (MIMO) with a Spatial-Multiplexing scheme is a topic of high interest for the next generation of wireless communications systems. In this paper, we propose to approach the Maximum Likelihood (ML) performance through the combination of a neighbourhood study and a Lattice Reduction (LR)-aided solution. Moreover, by introducing a neighbourhood study in the reduced domain, we propose in this paper a novel equivalent metric that is based on the combination of the LR-aided Minimum-Mean Square Error solution. We show that the proposed metric presents a relevant complexity reduction while maintaining near-ML performance. In particular, the corresponding computational complexity is polynomial in the number of antennas while it is shown to be independent of the constellation size. For a 4×4 MIMO system with 16-QAM modulation on each layer, the proposed solution is simultaneously near-ML and ten times less complex than the classical neighbourhood-based K-Best solution.*

Keywords: *MIMO, Sphere Decoder, Lattice-Reduction*

1. Introduction

Multiple-input multiple-output (MIMO) technology has gained a lot of attention in the last decade since it can improve link reliability without sacrificing bandwidth efficiency, or contrariwise it can improve bandwidth efficiency without losing link reliability. However, the main drawback of MIMO technology is the increased complexity of the detector when a Non-Orthogonal MIMO scheme is implemented. Indeed, although the performance of the Maximum Likelihood (ML) detector is symbol-wise optimal, its computational complexity increases exponentially with the number of transmit antennas and with the constellation size.

In the literature, some non-linear detectors have been introduced. The Sphere Decoder (SD) - based on a tree search-based neighbourhood study - is very popular due to its optimal performance [1]. However, this performance is reached at the detriment of an NP-hard complexity [1]. The authors of [2] have proposed a sub-optimal solution denoted as the *K*-Best solution, where *K* is the number of stored neighbours at each layer during the detection process. Aiming at reducing the neighbourhood size *K*, different solutions are proposed. For instance, the Sorted QR Decomposition (SQRD)-based dynamic *K*-Best (with a variable *K*), which leads to the famous SQRD-based Fixed Throughput SD (FTSD), is proposed in [3]. An alternative trend has been presented in the pioneering work of Wübben *et al.* in [4]. It consists in adding a pre-processing step, namely the Lattice Reduction (LR), aiming at applying a classical detection through a better-conditioned [4] channel matrix. This solution has been shown to offer the full reception diversity at the expense of a SNR offset in the detector's performance. However, this offset increases with large dimensional systems and high order modulations.

In this paper, we propose to combine the *K*-Best SD with the neighborhood size reduction through an efficient pre-processing step. This allows the SD process to apply a neighborhood study in a modified constellation domain. Then, we propose a novel modified ML equation based on the LR-Aided (LRA) Minimum-Mean Square Error (MMSE) detection. The proposed metric presents a large complexity reduction while maintaining near-ML performance. Moreover, its computational complexity is independent of the constellation size while it is polynomial with respect to the number of antennas. In particular, for a 4x4 MIMO system with 16-QAM modulation on each layer, the proposed solution is simultaneously near-ML and ten times less complex than the classical *K*-Best solution. We note that the complexity is fixed with such a detector, thus the exposed optimizations will induce a performance gain for a given neighborhood size or a reduction of the neighborhood size for a given Bit Error Rate (BER) target.

This paper is organized as follows. Section 2 presents the problem statement of the detection process. In section 3, we propose our generalized solution based on LR with the use of an efficient search centre and a reduced domain neighbourhood. In section 4, the performance of the presented detectors are provided, compared and discussed. Conclusions are drawn in section 5.

2.Problem Statement: Detection Process in the Original Domain

2.1. Sphere Decoder

It has been stated in the introduction that the *K*-Best solution approaches the optimal performance at the detriment of a large complexity, while the linear detectors present lower complexity with a strong penalty in performance. Hence, an optimal trade-off should be found.

Let us introduce an n_T-transmit and n_R-receive $n_T \times n_R$ MIMO system model. The received symbols vector could be then written as:

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{H}$ represents the $n_R \times n_T$ complex channel matrix that is assumed to be perfectly known at the receiver, $\boldsymbol{x}$ is the transmit symbols vector of dimension n_T where each entry is independently withdrawn from a constellation set ξ and $\boldsymbol{n}$ is an additive white Gaussian noise of dimension n_R and of variance σ^2. The basic idea of the SD to reach the optimal ML estimate $\widehat{\boldsymbol{x}}_{ML}$ while avoiding an exhaustive search is to examine only the lattice points that lie inside of a sphere having a radius d. The SD solution starts from the ML equation $\widehat{\boldsymbol{x}}_{ML} = \underset{x\in\xi^{n_T}}{\operatorname{argmin}} \|\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}\|^2$ and reads:

$$\widehat{\boldsymbol{x}}_{SD} = \underset{x\in\xi^{n_T}}{\operatorname{argmin}} \|\boldsymbol{Q}^H\boldsymbol{y} - \boldsymbol{R}\boldsymbol{x}\|^2 \leq d^2 \quad (2)$$

where $\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{R}$, with the classical QRD definitions [5].

The classical SD formula in (2) is centred on the received signal $\boldsymbol{y}$. The corresponding detector will be denoted in the following as the naïve SD. That is, in the case of a depth-first search algorithm [1], the first solution given by the algorithm is usually defined as the Babai point [6]. In our work, the definition is extended and the Babai point is denoted as the solution that is reached with no neighbourhood study. The QRD step in (2) actually aims at splitting the receiver computational complexity into a pre-processing stage (that depends on the channel only) and a processing step (that depends also on the receive data).

2.2. Lattice Reduction

Through the aforementioned considerations and by using the lattice definition in [6], the system model given in (1) rewrites, using the lattice reduction, as:

$$\boldsymbol{y} = \widetilde{\boldsymbol{H}}\boldsymbol{z} + \boldsymbol{n}, \quad (3)$$

where $\widetilde{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{T}$ and $\boldsymbol{z} = \boldsymbol{T}^{-1}\boldsymbol{x}$. The $n_T{\times}n_T$ complex matrix $\boldsymbol{T}$ is unimodular, namely its entries belong to the set $\mathbb{Z}_{\mathbb{C}} = \mathbb{Z} + j\mathbb{Z}$ of complex integers, with $j^2 = -1$, and $\boldsymbol{T}$ is such that $|\det\{\boldsymbol{T}\}| = 1$. The key idea of any LRA detection scheme is to understand that the finite set of transmitted symbols ξ^{n_T} used in (1) can be interpreted as a de-normalized, shifted then scaled version of the infinite set of complex integers subset $\subset \mathbb{Z}_{\mathbb{C}}^{n_T}$ used in (3), according to the relations offered in [6]. Using $\boldsymbol{T}$ in (3), we are now able to detect the transmitted symbols on different antennas in the modified domain. Indeed, the transmitted vector $\boldsymbol{x}$ could be deduced from z by using the relation $\boldsymbol{x} = \boldsymbol{T}\boldsymbol{z}$ [10].

The LRA detectors are very efficient in the sense of the high quality of their hard output, namely the ML diversity is reached within a constant SNR offset, while offering a low overall computational complexity. However, some drawbacks occur. In particular, the aforementioned SNR offset is important in the case of high order modulations, namely 16-QAM and 64-QAM, and with a large number of antennas. This issue is expected to be bypassed through an additional neighborhood study.

3.Proposed Detection Process in the Reduced Domain Neighbourhood

Joining the LRA and SD, called hereafter LRA-SD, by following the LR preprocessing step by any SD detector is not straightforward. The main issue lies in the consideration of the possibly transmit symbols vector in the reduced constellation since, unfortunately, the set of all possibly transmit symbols vectors cannot be predetermined. The reason for that is because the solution does not depend on the employed constellation only, but also on the $\boldsymbol{T}^{-1}$ matrix of (3) that mixes entries in $\boldsymbol{x}$. Consequently, the number of children in the tree search and their values are not known in advance. A brute-force solution is then to determine the set of all possibly transmit vectors in the reduced constellation, starting from the set of all possibly transmit vectors in the original constellation and then, switching to the reduced domain thanks to the $\boldsymbol{T}^{-1}$ matrix. Clearly, this operation must be done for each channel realization which is unfeasible in practice.

In the following, we will describe the preprocessing step and the reduction of the neighborhood size required for achieving quasi-ML estimation with a large complexity reduction.

3.1. Preprocessing

In [4], it has been shown the advantages of QRD in terms of performance when combined with ordering of the different SNR values (respectively SINR) on different antennas in the Zero-Forcing (ZF) (MMSE)-SQRD-based detectors. In this work, we propose to use the SQRD in the LRA-MMSE Extended (MMSEE) Ordered Successive Interference Cancellation (LRA-MMSEE-OSIC) detector.

In order to describe the MMSEE consideration, we introduce like in [4] an extended system model, namely the $(n_R + n_T) \times n_T$ matrix $\boldsymbol{H}_{ext}$ and the $(n_R + n_T)$ dimensional vector $\boldsymbol{y}_{ext}$ given by:

$$\mathbf{H}_{ext} = \begin{bmatrix}\mathbf{H}\\ \sigma\mathbf{I}\end{bmatrix} \text{ and } \mathbf{y}_{ext} = \begin{bmatrix}\mathbf{y}\\ \mathbf{0}\end{bmatrix} \quad (4)$$

The pre-processing step is similar to the ZF-SQRD given in [4] and the detection equals that of LRA-ZF SIC.

The LRA-MMSEE OSIC corresponds, to the best of the authors' knowledge, to the best pseudo-linear detector in the literature, in particular in the case of 4×4 MIMO systems with 4-QAM modulations on each layer [4]. The SQRD interest lies in the ordering of the detection symbols as a function of their S(I)NR and consequently it limits the error propagation in SIC procedures. Indeed from one side, it has been shown by Wübben *et al.* [4] that the optimum order offers a performance improvement, even if the ML diversity is not reached. On the other hand, it was shown that once the ML diversity is achieved through a LRA technique, the performance may be significantly improved with the solution in [4]. In order to deal with these statements, we introduce the Reduced Domain Neighbourhood (RDN) by using the following notations:

- $Q_{\xi^{n_T}}\{.\}$ is the quantization operator in the original domain constellation,
- $Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{.\}$ is the quantization operator in the reduced domain constellation,
- a is the power normalization and scaling coefficient (*i.e.* $2/\sqrt{2}$, $2/\sqrt{10}$, $2/\sqrt{42}$ for 4-QAM, 16-QAM and 64-QAM constellations, respectively)
- $\boldsymbol{d} = \frac{1}{2}\boldsymbol{T}^{-1}[1+j \quad \dots \quad 1+j]^T$ is a complex displacement vector.

In the literature, the classical LRA with a neighbourhood study is implicitly unconstrained LRA-ZF centred, which leads to a LRA-ZF SIC procedure with an RDN study at each layer. The exact formula has not been clearly provided but is implicitly given in [7] as:

$$\hat{\mathbf{z}}_{LRA-ZF\,SD} = \underset{\mathbf{z}\in\mathbb{Z}_{\mathbb{C}}^{n_T}}{\operatorname{argmin}} \left\| \widetilde{\mathbf{R}}(\mathbf{z}_{LRA-ZF} - \mathbf{z}) \right\|^2 \quad (5)$$

where $\widetilde{\mathbf{R}}$ is the well-known Lenstra-Lenstra-Lovász (LLL) algorithm based [8] LR algorithm output and $\mathbb{Z}_{\mathbb{C}}^{n_T}$ is the n_T-dimensional infinite set of complex integers.

3.2. Equivalent LRA-MMSE(E) centre

In the literature, and to the best of the author's knowledge, no convincing formula dealing with the LRA-MMSE(E) SD detection has been proposed. This is due to the non-factorable nature of the unconstrained LRA-MMSE(E) solution. In this work, we propose to do the factorization through the introduction of the following definition [9][12]:

***Definition**:* Two cost functions are *equivalent* iff :

$$\underset{\mathbf{x}\in\xi^{n_T}}{\operatorname{argmin}}\{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2\} = \underset{\mathbf{x}\in\xi^{n_T}}{\operatorname{argmin}}\{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2 + c\}, \quad (6)$$

where c is a constant. Obviously, the argument outputs are the same with both relations above.

Using (6), an unconstrained LRA-MMSE-centred solution can be derived from [12]. Under the assumption of using constant modulus constellations, solving the ML equation is equivalent to solve the following equation

$$\hat{\mathbf{z}}_{LRA-LD\,SD} = \underset{\mathbf{z}\in\mathbb{Z}_{\mathbb{C}}^{n_T}}{\operatorname{argmin}} \left\| \widetilde{\mathbf{R}}(\tilde{\mathbf{z}} - \mathbf{z}) \right\|^2, \quad (7)$$

where $\widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}} = \begin{cases} \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} & \text{in the ZF case,} \\ \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} + \sigma^2\mathbf{T}^H\mathbf{T} & \text{in the MMSE case,} \end{cases}$
$\tilde{\mathbf{z}}$ is any LRA (ZF or MMSE) unconstrained linear estimate and $\mathbf{T}$ is the transform matrix given in (3).

***Proof**:* Let us introduce any term c *s.t.* $\left\|\mathbf{y}-\widetilde{\mathbf{H}}\mathbf{z}\right\|^2 + c = \left\|\widetilde{\mathbf{R}}(\tilde{\mathbf{z}} - \mathbf{z})\right\|^2$, where $\tilde{\mathbf{z}}$ is a LRA-ZF (or LRA-MMSE) unconstrained linear estimate. This constant c is given by:

$$\begin{aligned}
c &= \left\|\widetilde{\mathbf{R}}(\tilde{\mathbf{z}} - \mathbf{z})\right\|^2 - \left\|\mathbf{y}-\widetilde{\mathbf{H}}\mathbf{z}\right\|^2 \\
&= (\tilde{\mathbf{z}} - \mathbf{z})^H\widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}}(\tilde{\mathbf{z}} - \mathbf{z}) - \left(\mathbf{y}-\widetilde{\mathbf{H}}\mathbf{z}\right)^H\left(\mathbf{y}-\widetilde{\mathbf{H}}\mathbf{z}\right) \\
&=^{(a)} \tilde{\mathbf{z}}^H\widetilde{\mathbf{G}}\tilde{\mathbf{z}} - \tilde{\mathbf{z}}^H\widetilde{\mathbf{G}}\mathbf{z} - \mathbf{z}^H\widetilde{\mathbf{G}}\tilde{\mathbf{z}} + \mathbf{z}^H\widetilde{\mathbf{G}}\mathbf{z} - \mathbf{y}^H\mathbf{y} + \\
&\quad \mathbf{y}^H\widetilde{\mathbf{H}}\mathbf{z} + \mathbf{z}^H\widetilde{\mathbf{H}}^H\mathbf{y} - \mathbf{z}^H\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\mathbf{z} \\
&=^{(b)} \mathbf{y}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{H}}^H\mathbf{y} - \mathbf{y}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{G}}\mathbf{z} - \\
&\quad \mathbf{z}\widetilde{\mathbf{G}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{H}}^H\mathbf{y} + \mathbf{z}^H\widetilde{\mathbf{G}}\mathbf{z} - \mathbf{y}^H\mathbf{y} + \mathbf{y}^H\widetilde{\mathbf{H}}\mathbf{z} + \\
&\quad \mathbf{z}^H\widetilde{\mathbf{H}}^H\mathbf{y} - \mathbf{z}^H\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\mathbf{z} \\
&=^{(c)} \mathbf{y}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{H}}^H\mathbf{y} + \mathbf{z}^H\left(\widetilde{\mathbf{G}} - \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}\right)\mathbf{z} - \mathbf{y}^H\mathbf{y}
\end{aligned}$$

where in (a), we substitute $\widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}}$ and $\widetilde{\mathbf{G}}$; in (b), we introduce $\tilde{\mathbf{z}} = \widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{H}}^H\mathbf{y}$ and $\tilde{\mathbf{z}}^H = \mathbf{y}^H\widetilde{\mathbf{H}}\widetilde{\mathbf{G}}^{-1}$ and, in (c):

$$\widetilde{\mathbf{G}} = \widetilde{\mathbf{R}}^H\widetilde{\mathbf{R}} = \begin{cases} \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} & \text{in the ZF case,} \\ \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} + \sigma^2\mathbf{T}^H\mathbf{T} & \text{in the MMSE case.} \end{cases} \quad (8)$$

In the ZF case, $\widetilde{\mathbf{H}}\widetilde{\mathbf{G}}^{-1}\widetilde{\mathbf{H}}^H = \widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}(\widetilde{\mathbf{H}}^H)^{-1}\widetilde{\mathbf{H}}^H = \mathbf{I}$ and $\widetilde{\mathbf{G}} - \widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} = \mathbf{0}$. Consequently $c = 0$.
In the MMSE case, $c = \mathbf{y}^H\left[\widetilde{\mathbf{H}}(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} + \sigma^2\mathbf{T}^H\mathbf{T})^{-1}\widetilde{\mathbf{H}}^H - \mathbf{I}\right]\mathbf{y} + \sigma^2\mathbf{z}^H\mathbf{T}^H\mathbf{T}\mathbf{z}$ which is a constant term in $\mathbf{x}$ iff the signal $\mathbf{x}$ entries are of constant modulus since $\sigma^2\mathbf{z}^H\mathbf{T}^H\mathbf{T}\mathbf{z} = \sigma^2\mathbf{x}^H\mathbf{x}$.■

In the proof above, the use of constant modulus constellations for $\mathbf{x}$ entries is not limiting because, first, we can consider that this assumption is respected in mean for a large number of transmitting antennas n_T; second, any *M*-QAM constellation can be considered as a linear sum of 4-QAM points [12].
The formula introduced in (7) offers an equivalent metric, in the reduced domain, to the metric introduced in [1],[2],[3]. The main difference however relies on the neighbourhood study nature. In the case of an RDN study, the equivalent channel matrix $\widetilde{\mathbf{H}}$ given in (7) is considered and is noticed to be almost - while not exactly - orthogonal. Consequently, the independent detection layer by layer of the symbols vector $\mathbf{x}$ does not exactly correspond to its joint detection since the mutual influence of the transformed signal $\mathbf{z}$ is still present, thus it highlights the advantage of a neighbourhood study. It should be noticed that an RDN study will require a smaller size compared to an Original Domain Neighbourhood (ODN) study.

3.3. Reduced-Domain Neighbourhood study

The main issue of the RDN study lies in the generation of the set of possibly transmit symbols, at the considered layer. Contrary to classical SD, it cannot be pre-determined since the solution does not depend on the employed constellation only, but also on the $\mathbf{T}^{-1}$ matrix in (3) which mixes the layers. Thus, their exact pre-determination may be only done jointly.

During the exploration of possible solutions, two major drawbacks arise. First, the geometry of constellations in the reduced domain may be rather complex and even induce non-adjacent symbols. In particular, some candidates may not map to any existing constellation points in the original domain [7]. Second, the reduced domain does not consider any boundary region as in the case of QAM constellations. Thus, it induces no limitation in the neighbourhood size. No satisfying brute-force solution exists, as stated in Section 2.

In order to efficiently solve the neighborhood generation issue recalled above, we propose in this work to use Schnorr-Euchner (SE) [11] enumeration. Starting from the LRA principle by passing from $\mathbf{x}$ to $\mathbf{z}$ through (3), a neighbourhood centred on the LRA-MMSE(E) OSIC (before quantization) solution, *i.e.* $\tilde{\mathbf{z}}_k$, is considered at each layer k. The RDN generation is processed for a bounded number of *N* possibilities in a SE fashion, namely according to an increasing Partial Euclidian Distances (PED) [2] from $\tilde{\mathbf{z}}_k$ at each layer, such that:

$$\begin{aligned}
&\mathbf{z}_k = Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\}, \\
&Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} + 1, Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} + j,\ Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} - 1, Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} - j, \quad (9) \\
&Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} + 2, Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} + 2j, Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} - 2, Q_{\mathbb{Z}_{\mathbb{C}}^{n_T}}\{\tilde{\mathbf{z}}_k\} - 2j, \ldots
\end{aligned}$$

The SE strategy aims at finding the correct decision as early as possible, leading to a safe early termination criterion. In the proposed technique, the *K* best solutions are stored at each layer. Thus it is denoted as RDN LRA-MMSE(E) *K*-Best detector.

3.4. Block Diagram of the proposed technique

The general principle of any LRA detector is depicted in a block-diagram manner in Figure 1. The mapping of any estimate $\hat{\mathbf{z}}$ (or list of estimates) from the reduced domain to the estimate $\tilde{\mathbf{x}}$ in the original domain is processed through the $\mathbf{T}$ matrix multiplication (see Equation (3)). The additional quantization step aims at removing duplicate symbols vectors outputs in the case of a list of solutions. Thus it should be noted that it is useful in the case of a neighbourhood study only.

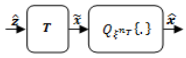

Figure 1- Block-diagram of any LRA procedure

The search is achieved independently of the search centre calculation. In particular, the detector's computational complexity is split between the search centre calculation step and the neighbourhood search step as depicted in Figure 2.

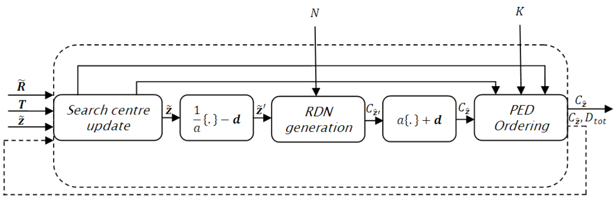

Figure 2- Block-diagram of any RDN LRA *K*-Best procedure

By considering the k-th layer and the knowledge of the $C_{\hat{\mathbf{z}}_{k+1:n_T}}$ symbols estimated at the layers *k+1* up to n_T, the $\tilde{\mathbf{z}}_k$ unconstrained point can be provided and then has to be de-normalized and shifted, thus making $\tilde{\mathbf{z}}'_k$ to belong to $\mathbb{Z}_{\mathbb{C}}^{n_T}$. At this step, the RDN study addressed in Section 3.3 arises. After quantization, de-shifting and normalization, the $C_{\hat{\mathbf{z}}_k}$ estimates at the k-th layer are obtained. That is, all the PEDs calculation and ordering (detailed in [2]) are applied explicitly. The K best solutions, namely with the lowest Cumulative Euclidean Distances (CED) (D_{tot}) in the reduced domain, are stored ($C_{\hat{\mathbf{z}}}$).
The search centre is updated at each layer, resulting in an improved performance. In particular, the next (*k-1*)-th layer is considered, until the whole list of symbols vector is detected.
By now adding, prior to the RDN study, the pre-processing steps, namely the SQRD-based LR technique, and the computation of a close-to-ML linear unconstrained estimate using the LRA-MMSEE a complete description of the detection structure can be obtained. A final step of the detector should be added to the detection process. Indeed, the list of possible symbols output has to be re-ordered according to the ML metrics in the original domain and duplicate solutions should be removed. This is due to the presence of noise that makes some candidates to be mapped on non-legitimate constellation points in the reduced constellation, leading to inappropriate points in the original constellation. The symbols vector associated to the minimal metric becomes the hard decision output of the detector and offers a near-ML solution.

4.System Performance

In this section, we present the system performance of the aforementioned reference solutions and we compare them to our proposed RDN LRA-MMSEE *K*-Best solution.
Figure 3, Figure 4 and Figure 5 depict the uncoded BER for 4×4 MIMO scheme using 16-QAM modulations with *K*=2, *K*=4 and *K*=16, respectively. The RDN size is set to *N*=5, which is shown in the simulations to be sufficient for our detector (not for the reference LRA-ZF *K*-Best). Some notable points have to be highlighted from these figures. As previously claimed and contrary to the RDN LRA *K*-Best, the ODN *K*-Best solutions (ZF and MMSE) do not reach the ML diversity for a reasonable neighbourhood size, even if there is a decrease of the SNR offset in the ODN MMSE *K*-Best case. However, the classical RDN LRA-ZF *K*-Best technique presents a BER offset in the low SNR range, due to errors propagation. Consequently, there exists a switching point from low to high SNR between LRA detectors and others. This aspect is removed through the use of the SQRD in the RDN LRA-MMSEE *K*-Best solution. The full diversity is still reached and the BER offset in low SNR regime is highly reduced compared to the RDN LRA-ZF *K*-Best solution. Even for a small RDN size (*K*=2) and for 16-QAM constellations, its performance is now close-to-ML as shown in Figure 3.

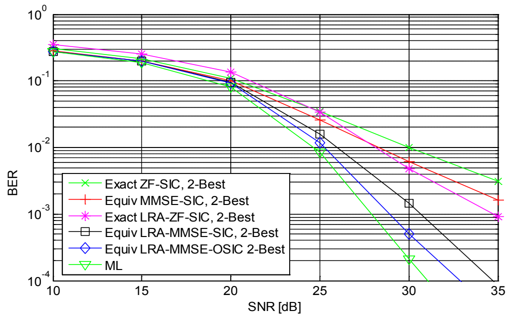

Figure 3- Uncoded BER, 4×4 complex Rayleigh channel, 16-QAM modulation on each layer, *K*=2

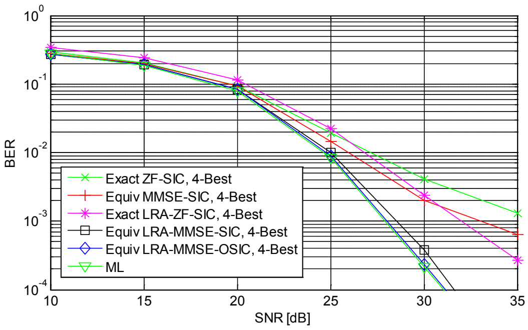

Figure 4- Uncoded BER, 4×4 complex Rayleigh channel, 16-QAM modulation on each layer, *K*=4

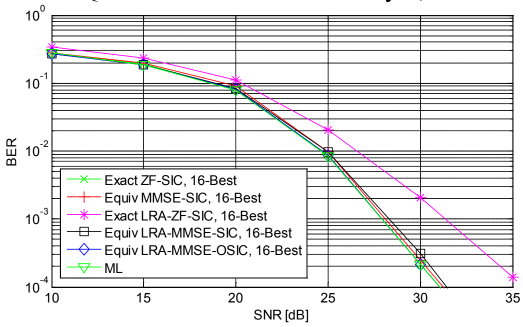

Figure 5- Uncoded BER, 4×4 complex Rayleigh channel, 16-QAM modulation on each layer, *K*=16

The SNR losses compared to ML solution of all the plotted techniques are given in Table 1. They have been measured for an uncoded BER of 10^{-4}. Also, the corresponding computational complexities are given in Table 2 in terms of real multiplications (MUL) operations and for an RDN size $N = 5$. Now, comparing the LRA-MMSE(E) solution with the *K*-Best solution, it is straightforward to observe that the proposed technique requires less neighbours (*K*=2) while the *K*-Best requires *K*=16. Even if the proposed LRA-MMSE(E) solution is two times more complex with 4-QAM constellations, it offers near-ML performance and in particular an SNR gain of 0.3 dB at a BER of 10^{-4}. The most interesting point concerns higher order modulations: for the 16-QAM modulation, the estimated complexity of the proposed solution is ten times less complex than the classical one, for the same performance result (see underlined results in Table 2 for comparison purposes).

Due to space limitation, the 16-QAM case only has been presented in this section while near-ML performance is achieved anyway for 64-QAM. Nevertheless, we should mention an important point. Benefits of LRA SD technique are limited in the case of the widely used 4-QAM constellation, due to the existence of an implicit constraint from the 4-QAM constellation construction. In particular, the quantization operation $Q_{\xi^{n_T}}\{.\}$ induces a constraint that eliminates nearby lattice points that do not belong to ξ^{n_T}. This aspect destroys a part of the LRA benefit and cannot be corrected despite the increase of the neighbourhood study size. Indeed, many lattice points considered in the RDN would be associated with the same constellation point after quantization in the original constellation. In the case of larger constellation orders, the LRA benefit is more effective, as depicted in Table 1.

Finally, we should note that the proposed RDN LRA-MMSEE *K*-Best solution is particularly efficient in the case of ill-conditioned MIMO systems, *i.e.* spatially correlated antennas systems, due to the LR step and in the case of high order constellations size due to the neighbourhood study (two LTE-A norm requirements [13]).

Table 1- SNR loss compared to the ML solution

	SNR loss (4-QAM)				**SNR loss (16-QAM)**			
Technique	*K*=1	*K*=2	*K*=3	*K*=4	*K*=1	*K*=2	*K*=4	*K*=16
ZF K-Best	>7.6	>7.6	>7.6	0.36	>5.0	>5.0	>5.0	0.00
MMSE K-Best	>7.6	>7.6	6.21	0.30	>5.0	>5.0	>5.0	0.09
LRA-ZF K-Best	4.43	2.90	1.92	1.71	3.21	2.04	1.27	0.62
LRA-MMSE K-Best	2.90	0.73	0.52	0.27	2.12	0.76	0.53	0.40
LRA-MMSEE K-Best	0.80	0.01	0.00	0.00	1.62	0.02	0	0.00

Table 2- Computational complexities in *MUL*

	MUL (4-QAM)				**MUL (16-QAM)**			
Technique	*K=1*	*K=2*	*K=3*	*K=4*	*K=1*	*K=2*	*K=4*	*K=16*
ZF K-Best	*156*	*300*	*444*	*588*	*624*	*1200*	*2352*	*9264*
MMSE K-Best	*156*	*300*	*444*	*588*	*624*	*1200*	*2352*	*9264*
LRA-ZF K-Best	*510*	*946*	*1382*	*1818*	*510*	*946*	*1818*	*2254*
LRA-MMSE K-Best	*510*	*946*	*1382*	*1818*	*510*	*946*	*1818*	*2254*
LRA-MMSEE K-Best	*510*	*946*	*1382*	*1818*	*510*	*946*	*1818*	*2254*
ML	*16384*				*4194304*			

5.Conclusion

In this paper, the RDN LRA-MMSE(E) SD has been proposed, with a *K*-Best neighbourhood generation. Detailed computational complexity estimation has been provided and compared to the state of the art. In particular, we have shown that the proposed detector outperforms existing solutions while requiring much lower complexity. Interestingly, the corresponding computational complexity is independent of the constellation size and polynomial in the number of antennas, again while reaching the ML performance. It implies 10 times lower computational complexity compared to the classical *K*-Best for a 4×4 MIMO system, with 16-QAM modulation on each layer and still better for higher-order constellations.

Acknowledgements

This work was supported by ST-ERICSSON and by the European Commission under the SYMPA project.

References

[1] B. Hassibi, and H. Vikalo, "On the expected complexity of sphere decoding," *Signal, Systems and Computers, Asimolar Conference on*, pp. 1051–1055, 2001.

[2] K.-W. Wong, C.-Y. Tsui, S.-K. Cheng, W.-H. Mow, "A VLSI Architecture of a *K*-Best Lattice Decoding Algorithm For MIMO Channels," C*ircuits and Systems, IEEE International Symposium on*, vol. 3 pp. 273.276, May 2002.

[3] L. G. Barbero, and J. S. Thompson, "A fixed-complexity MIMO detector based on the complex sphere decoder," *Signal Processing Advances for Wireless Communications, Workshop on*, pp.1–5, 2006.

[4] D. Wüebben, R. Böhnke, V. Kühn, and K.-D. Kammeyer, "MMSE-based lattice-reduction for near-ML detection of MIMO systems," *Smart Antennas, ITG Workshop on*, pp. 106–113, 2004.

[5] S. Aubert, and M. Mohaisen, "From Linear Equalization to Lattice-Reduction-Aided Sphere-Detector as an answer to the MIMO Detection problematic in Spatial Multiplexing Systems," *Vehicular Technologies*, INTECH, Feb. 2011.

[6] E. Agrell, T. Eriksson, E. Vardy, and K. Zeger, "Closest Point Search in Lattices," *Information Theory, IEEE Transactions on*, vol. 48, no 8, pp: 2201-2214, 2002.

[7] X.-F. Qi, and K. Holt, "A Lattice-Reduction-Aided Soft Demapper for High-Rate Coded MIMO-OFDM Systems," *Signal Processing Letters, IEEE*, vol. 14, no. 5, pp. 305-308, May 2007.

[8] A.K. Lenstra, H.W. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, no. 4, pp. 515–534, 1982.

[9] L. Wang, L. Xu, S. Chen, and L. Hanzo, "MMSE Soft-Interference-Cancellation Aided Iterative Center-Shifting *K*-Best Sphere Detection for MIMO Channels," *Communications, IEEE International Conference on*, pp.3819-3823, May 2008.

[10] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "Communication Over MIMO Broadcast Channels Using Lattice-Basis Reduction", IEEE Trans. On Information Theory, Vol. 53, No. 12, Dec. 2007

[11] C. Schnorr, M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Mathematical Programming*, pp.:181-199, Sept. 1994

[12] T. Cui, and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," *Communications Letters, IEEE*, vol. 9, no. 5, pp. 423-425, May 2005.

[13] E.-U. Technical Specification Group RAN, "36.101 User Equipment (UE) radio transmission and reception v8.8.0," Tech. Rep., Sept. 2009.