Sphere Decoding for Spatial Modulation

Abdelhamid Younis, Marco Di Renzo, Raed Mesleh, Harald Haas

Sphere Decoding for Spatial Modulation

Abdelhamid Younis[1], Marco Di Renzo[2], Raed Mesleh[3], and Harald Haas[1]

[1]*The University of Edinburgh, College of Science and Engineering*
Institute for Digital Communications, Joint Research Institute for Signal and Image Processing
King's Buildings, Mayfield Road, Edinburgh, EH9 3JL, UK
[2]*L2S, UMR 8506 CNRS – SUPELEC – Univ Paris–Sud*
Laboratory of Signals and Systems (L2S), French National Center for Scientific Research (CNRS)
École Supérieure d'Électricité (SUPÉLEC), University of Paris–Sud XI (UPS)
3 rue Joliot–Curie, 91192 Gif–sur–Yvette (Paris), France
[3]*Jacobs University Bremen, School of Engineering and Science*
Campus Ring 1, Research 1, 28759 Bremen, Germany
E–Mail: {a.younis, h.haas} @ed.ac.uk, marco.direnzo@lss.supelec.fr, r.mesleh@jacobs–university.de

Abstract—**In this paper, Sphere Decoding (SD) algorithms for Spatial Modulation (SM) are developed to reduce the computational complexity of Maximum–Likelihood (ML–) optimum detectors, which foresee an exhaustive search of the whole search space and have a complexity that linearly increases with the product of number of transmit–antenna, receive–antenna, and size of the modulation scheme. Three SDs specifically designed for SM are proposed and analyzed in terms of Bit Error Probability (BEP) and computational complexity. By judiciously choosing some key parameters, *e.g.*, the radius of the sphere centered around the received signal, it is shown that the proposed algorithms offer the same BEP as ML–optimum detection, with a significant reduction of the computational complexity. Also, it is shown that none of the proposed SDs is always superior to the others, but the best SD to use depends on the system setup, *i.e.*, the number of transmit–antenna, receive–antenna, and the size of the modulation scheme. The computational complexity trade–off offered by the proposed solutions is studied via analysis and simulation, and numerical results are shown to validate our findings.**

Index Terms—**Multiple–Input–Multiple–Output (MIMO) Systems, Spatial Modulation (SM), Sphere Decoding (SD).**

I. INTRODUCTION

Spatial Modulation (SM) is a recently proposed transmission technology for Multiple–Input–Multiple–Output (MIMO) wireless systems, which has been originally proposed to increase the spectral efficiency of single–antenna systems by avoiding Inter–Channel Interference (ICI) [1]. This is attained through the adoption of a new modulation and coding scheme, which foresees: i) the activation, at each time instance, of a single antenna that transmits a given data symbol (*constellation symbol*), and ii) the exploitation of the spatial position (index) of the active antenna as an additional dimension for data transmission (*spatial symbol*). Both *constellation symbol* and *spatial symbol* depend on the incoming data bits. Thereby, an overall increase, by the base–two logarithm of the number of transmit–antenna, of the spectral efficiency is achieved, while still retaining a complexity comparable to single–antenna systems.

In particular, at the receiver the Maximum Likelihood (ML) optimum decoder is a simple single–stream detector, which performs an exhaustive search over the whole *constellation symbol* and *spatial symbol* space, and whose computational complexity ($\mathcal{C}$) linearly increases with the product of transmit–antenna (N_t), receive–antenna (N_r), and size of the modulation scheme (M), *i.e.*, $\mathcal{C} \propto MN_tN_r$ [2]. Unlike other spatial multiplexing schemes for MIMO systems, such as the V–BLAST (Vertical Bell Laboratories Layered Space–Time) [3], [4], there is a substantial reduction in receiver complexity, as no multi–stream detectors with exponential–growing complexity with N_t are required. In addition to this significant complexity reduction, SM also outperforms many conventional single–antenna and multi–antenna wireless systems [1], [5], thus potentially being an appealing transmission concept for the next generation of wireless systems.

In spite of its low–complexity implementation and superior performance results, there still is potential for further computational complexity reductions, especially when: i) high spectral efficiencies are required (*i.e.*, the product MN_t is large), or ii) high diversity gains and, thus, low error probabilities, are needed (*i.e.*, N_r is large). Furthermore, complexity issues become even more pressing when both MN_t and N_r are large. Motivated by these considerations, some recent research works have focused on developing low–complexity detectors for SM. For example, in [1] and [6] two sub–optimal two–step detectors based on heuristics are proposed. However, in [2] it is shown that the detector in [1] belongs to the family of non–exact methods [7], and is, in general, a few dB worse than ML–optimum detection. On the other hand, in [8] an exact low–complexity detector for SM has been proposed, which is based on the Sphere Decoding (SD) algorithm [9]. Therein, it is shown that the proposed solution has a computational complexity that is bounded by $8MN_t \leq \mathcal{C} \leq 8MN_tN_r$, and provides error performance very close to the ML–optimum detector. This SD–based detector is especially suitable when the number of receive–antenna N_r is very large, as it reduces the size of the search space related to the multiple antennas at the receiver (we denote this search space as "receive search space"). However, it has two main limitations: i) it does not reduce the dimension of the search space due to the number, N_t, of transmit–antenna and the size, M, of the signal modulation diagram (we denote this search space as "transmit search space"), which prevents the detector to achieve a significant reduction in computational complexity

when high data rates are required (*i.e.*, when both N_t and M are large), and ii) the detector has the same complexity as the ML–optimum decoder when $N_r = 1$. In general, the reduction in decoding complexity is not very high when N_r is small, as often happens in the downlink of cellular systems. The detector in [8] is here called Receiver–centric SD (Rx–SD).

Motivated by these considerations, in this paper we move from the results in [8] and propose two new low–complexity detectors for SM, which are based on the SD principle. The first solution aims at reducing the "transmit search space", and, thus, can be seen as a complementary solution to [8]. In particular, as opposed to [8], the proposed decoder is suitable when either N_t or M, or both N_t and M are large. On the other hand, it keeps the "receive search space" the same as the original ML–optimum decoder. This detector is called Transmitter–centric SD (Tx–SD). On the other hand, the second solution combines both Rx–SD and Tx–SD with the aim of reducing the complexity of the ML–optimum receiver in both the receive and transmit search spaces. This detector is called Combined–SD (C–SD). More specifically, the Tx–SD detector is based on a simplified implementation of the conventional SD proposed in [7], which exploits the peculiar property of SM that only a single antenna is active at any time instance. Due to space constraints, in this paper we focus on the so–called non–underdetermined MIMO setup with $N_t \le N_r$. In the recent period, some efficient SD methods for the underdetermined MIMO setup have been proposed (see, *e.g.*, [10], [11], [12] and references therein). However, the analysis of this setup for SM is postponed to a future research contribution.

In this paper, we provide a careful study of the performance of these three detectors, along with an accurate comparison of their computational complexity. Numerical results show that the proposed solutions provide a substantial reduction in computational complexity with respect to the ML–optimum detector, and no loss in the Bit Error Probability (BEP). Furthermore, it is shown that the Rx–SD is less complex than the C–SD when M is not very large, while the C–SD is the best solution when either M is large or N_r is small. The reason why Rx–SD outperforms C–SD when M is not very large is due to the some pre–computations required by the SD algorithm to determine the actual points of the reduced space to search through.

The reminder of this paper is organized as follow. In Section II, the system model along with the ML–optimum and Rx–SD detectors are summarized. In Section III, the new Tx–SD and C–SD receivers are described. In Section IV, an accurate analysis of the computational complexity of Tx–SD and C–SD is performed. In Section V, some numerical results are shown to compare the proposed receivers. Finally, Section VI concludes the paper.

II. SYSTEM MODEL

A. SM Modulator

SM works as follows [1]. The bitstream emitted by a binary source is divided into blocks containing $\log_2(N_t) + \log_2(M)$ bits each, with $\log_2(N_t)$ and $\log_2(M)$ being the number of bits needed to identify the *spatial symbol* and the *constellation*

symbol, respectively. Each block is split into two sub–blocks of $\log_2(N_t)$ and $\log_2(M)$ bits each, and the following mapping rule is used:

- The bits in the first sub–block are used to select the antenna that is switched on for data transmission, while all the other transmit–antenna are kept silent. In this paper, the actual transmit–antenna that is active for transmission is denoted by ℓ_t, with $\ell_t \in \{1, 2, \ldots, N_t\}$.
- The bits in the second sub–block are used to choose a symbol in the signal–constellation diagram. Without loss of generality, Quadrature Amplitude Modulation (QAM) is considered. In this paper, the actual complex symbol emitted by the transmit–antenna ℓ_t is denoted by s_t, with $s_t \in \{s_1, s_2, \ldots, s_M\}$.

Accordingly, the $N_t \times 1$ transmitted vector is:

$$\mathbf{x}_{\ell_t, s_t} = \left[\mathbf{0}_{1 \times (\ell_t - 1)}, s_t, \mathbf{0}_{1 \times (N_t - \ell_t)} \right]^T \tag{1}$$

where $(\cdot)^T$ denotes transpose operation, and $\mathbf{0}_{p \times q}$ is a $p \times q$ matrix with all–zero entries.

B. Channel Model

The modulated vector, $\mathbf{x}_{\ell_t, s_t}$, in (1) is transmitted through a frequency–flat $N_r \times N_t$ MIMO fading channel with impulse response $\mathbf{H}$. In this paper, a Rayleigh fading channel model is assumed, and, thus, the entries of $\mathbf{H}$ are modeled as complex independent and identically distributed (i.i.d.) Gaussian random variables with zero–mean and unit–variance.

Thus, the $N_r \times 1$ received vector can be written as follows:

$$\mathbf{y} = \mathbf{H} \mathbf{x}_{\ell_t, s_t} + \mathbf{n} \tag{2}$$

where $\mathbf{n}$ is the N_r–dimensional Additive White Gaussian Noise (AWGN) with zero–mean and variance σ^2 per dimension at the receiver input.

From (1), (2) simplifies as follows:

$$\mathbf{y} = \mathbf{h}_{\ell_t} s_t + \mathbf{n} \tag{3}$$

where $\mathbf{h}_{\ell_t}$ is the ℓ_t–th column of $\mathbf{H}$.

C. ML–Optimum Detector

The optimum detector based on the ML principle has been derived in [2]:

$$
\begin{aligned}
\left[\hat{\ell}_t^{(\mathrm{ML})}, \hat{s}_t^{(\mathrm{ML})} \right] &= \underset{\substack{\ell \in \{1, 2, \ldots N_t\} \\ s \in \{s_1, s_2, \ldots s_M\}}}{\arg \min} \left\{ \| \mathbf{y} - \mathbf{h}_\ell s \|_{\mathrm{F}}^2 \right\} \\
&= \underset{\substack{\ell \in \{1, 2, \ldots N_t\} \\ s \in \{s_1, s_2, \ldots s_M\}}}{\arg \min} \left\{ \sum_{r=1}^{N_r} |y_r - h_{\ell, r} s|^2 \right\}
\end{aligned}
\tag{4}
$$

where $\| \cdot \|_{\mathrm{F}}^2$ is the Frobenius norm, $\hat{}$ identifies the estimated spatial and constellation symbols, and y_r and $h_{\ell, r}$ is the r–th entry of $\mathbf{y}$ and $\mathbf{h}_\ell$, respectively.

The computational complexity of (4), in terms of real multiplications, is equal to:

$$\mathcal{C}_{\mathrm{ML}} = 8 M N_t N_r \tag{5}$$

as the ML detector searches through the whole transmit and receive search spaces.

D. Rx–SD Detector

In [8], a reduced–complexity and close–to–optimal BEP–achieving decoder is proposed, which, as mentioned in Section I, aims at reducing the receive search space. The detector can formally be written as follows:

$$\left[\hat{\ell}_t^{(\text{Rx–SD})}, \hat{s}_t^{(\text{Rx–SD})}\right] = \underset{\substack{\ell \in \{1,2,\dots N_t\} \\ s \in \{s_1, s_2, \dots s_M\}}}{\arg\min} \left\{ \sum_{r=1}^{\tilde{N}_r(\ell,s)} |y_r - h_{\ell,r}s|^2 \right\} \tag{6}$$

where $1 \leq \tilde{N}_r(\ell, s) \leq N_r$ is the number of correlations in the receive search space that is computed by using the SD principle. Note that $\tilde{N}_r(\cdot, \cdot)$ can be different for each point in the transmit search space, i.e., for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$.

More specifically, the Rx–SD receiver works as follows. Given a sphere with radius R, the receiver computes the set of optimal $\tilde{N}_r(\ell, s)$ for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$ as follows:

$$\tilde{N}_r(\ell, s) = \min_{n \in \{1,2,\dots N_r\}} \left\{ n \left| \sum_{r=1}^n |y_r - h_{\ell,r}s|^2 > R^2 \right. \right\} \tag{7}$$

In other words, for each $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$, the Rx–SD receiver does not combine, according to the Maximal Ratio Combining (MRC) principle [2], the signals received by all the N_r antennas at the receiver, but it keeps combining the received signals until the Euclidean norm in (6) gives a point that lie inside a sphere of radius R and centered around the received signal itself. Since the number of antennas, $\tilde{N}_r(\cdot, \cdot)$, that are actually used for MRC are computed adaptively for each point in the transmit search space, in [8] it is shown that there is no loss in either diversity or coding gain: the BEP is exactly the same as the ML detector in (4).

In [8], the interested reader can find an accurate analysis of the computational complexity of the Rx–SD detector along with an efficient method to choose the radius R, which significantly affects the performance of the algorithm. In this paper, we simply mention that:

- The complexity of the Rx–SD receiver is given by:

$$\mathcal{C}_{\text{Rx–SD}} = 8 \sum_{\ell=1}^{N_t} \sum_{s=1}^{M} \tilde{N}_r(\ell, s) \tag{8}$$

It is easy to show that $\mathcal{C}_{\text{Rx–SD}}$ lies in the interval $8MN_t \leq \mathcal{C}_{\text{Rx–SD}} \leq 8MN_tN_r$, where the lower bound corresponds to the scenario where $\tilde{N}_r(\ell, s) = 1$ for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$, and the upper bound corresponds to the scenario where $\tilde{N}_r(\ell, s) = N_r$ for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$. In other words, in the best–case scenario Rx–SD has the same complexity as a Multiple–Input–Single–Output (MISO) system, while in the worst-case scenario it has the same complexity as the ML–optimum detector in (4). These results suggest that the larger N_r, the higher the potential gain with respect to the ML–optimum receiver. Let us note that the Rx–SD solution has no pre–computations with respect to the ML–optimum detector. In fact,

$\tilde{N}_r(\ell, s)$ for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$ in (7) are implicitly computed when solving the hypothesis–detection problem in (6).

- The radius R can be chosen as $R = 2\alpha N_r \sigma^2$, where α is the solution of the equation as follows [7], [8]:

$$\frac{\gamma(N_r, 2\alpha N_r)}{\Gamma(N_r)} = 1 - \varepsilon \tag{9}$$

and $\gamma(x, a) = \int_0^a \xi^{x-1} \exp(-\xi) \, d\xi$ is the lower incomplete Gamma function, $\Gamma(x) = \int_0^{+\infty} \xi^{x-1} \exp(-\xi) \, d\xi$ is the Gamma function, and ε is an arbitrary small value close to zero, e.g., $\varepsilon = 10^{-6}$.

III. New Low–Complexity Sphere Decoders for SM

As anticipated in Section I, in this section two new SDs for SM are introduced. The first one (Tx–SD), aims at reducing the transmit search space, as opposed to the Rx-SD search space, and the second one (C–SD) combines Rx–SD and Tx–SD decoders in order to reduce both transmit and receive search spaces.

A. Tx–SD Detector

The Tx–SD for SM is a modified version of the well–known SD for MIMO systems [7], [9]. It, however, exploits the peculiar property of SM that only a single antenna is active for transmission. More specifically, similar to conventional SDs, the Tx–SD scheme reduces the number of points (ℓ, s) for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$ to be searched through in (4), i.e., the transmit search space, by computing the Euclidean distances only for those points that lie inside a sphere with radius R and centered around the received signal. However, unlike conventional SDs, in our scheme the set of points inside the sphere are much simpler to be computed, as there is only a single active antenna in SM. In this section, we show how to compute these intervals. While in the next section, we provide an estimate of the computational complexity of the proposed detector.

The analytical derivation follows the notation in [7], which here is briefly summarized to make the paper self–contained. First, for ease of analytical derivation, the complex–valued model in (2) is replaced by its real–valued equivalent, as follows [13]:

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{x}}_{\ell_t, s_t} + \bar{\mathbf{n}} \tag{10}$$

where $\bar{\mathbf{y}}$ is a $2N_r \times 1$ vector, $\bar{\mathbf{H}}$ is a $2N_r \times 2N_t$ matrix, $\bar{\mathbf{x}}_{\ell_t, s_t}$ is a $2N_t \times 1$ vector, and $\bar{\mathbf{n}}$ is a $2N_r \times 1$ vector defined as follows:

$$\begin{cases} \bar{\mathbf{y}} = \left[\text{Re}\{\mathbf{y}^T\}, \text{Im}\{\mathbf{y}^T\}\right]^T \\ \bar{\mathbf{H}} = \begin{bmatrix} \text{Re}\{\mathbf{H}\} & \text{Im}\{\mathbf{H}\} \\ -\text{Im}\{\mathbf{H}\} & \text{Re}\{\mathbf{H}\} \end{bmatrix} \\ \bar{\mathbf{x}}_{\ell_t, s_t} = \left[\text{Re}\{\mathbf{x}_{\ell_t, s_t}^T\}, \text{Im}\{\mathbf{x}_{\ell_t, s_t}^T\}\right]^T \\ \bar{\mathbf{n}} = \left[\text{Re}\{\mathbf{n}^T\}, \text{Im}\{\mathbf{n}^T\}\right]^T \end{cases} \tag{11}$$

and $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ denote real and imaginary parts, respectively.

By performing QR factorization of the matrix $\bar{\mathbf{H}}$, (10) can

$$\Theta_R = \left\{ (\ell, s) \text{ with } \ell \in \{1, 2, \dots N_t\} \text{ and } s \in \{s_1, s_2, \dots s_M\} \big| \left\| \bar{\mathbf{y}} - \bar{\mathbf{H}} \bar{\mathbf{x}}_{\ell,s} \right\|_F^2 \leq R^2 \right\}$$

$$= \left\{ (\ell, s) \text{ with } \ell \in \{1, 2, \dots N_t\} \text{ and } s \in \{s_1, s_2, \dots s_M\} \big| \sum_{i=1}^{2N_t} \left(\bar{z}_i - \sum_{j=i}^{2N_t} \bar{p}_{i,j} \bar{x}_j (\ell, s) \right)^2 \leq R_Q^2 \right\} \quad (14)$$

be re–written as follows [7]:

$$\bar{\mathbf{y}} = \bar{\mathbf{Q}} \begin{bmatrix} \bar{\mathbf{P}} \\ \mathbf{0}_{(2N_r - 2N_t) \times 2N_t} \end{bmatrix} \bar{\mathbf{x}}_{\ell_t, s_t} + \bar{\mathbf{n}} \quad (12)$$

where $\bar{\mathbf{P}}$ is a $2N_t \times 2N_t$ upper triangular matrix, $\bar{\mathbf{Q}} = \begin{bmatrix} \bar{\mathbf{Q}}_1 & \bar{\mathbf{Q}}_2 \end{bmatrix}$, and $\bar{\mathbf{Q}}_1$, $\bar{\mathbf{Q}}_2$ are $2N_r \times 2N_t$ and $2N_r \times (2N_r - 2N_t)$ matrices, respectively.

The Tx–SD scheme can be formally written as follows:

$$\left[\hat{\ell}_t^{(\text{Tx-SD})}, \hat{s}_t^{(\text{Tx-SD})} \right] = \arg\min_{(\ell,s) \in \Theta_R} \left\{ \left\| \bar{\mathbf{y}} - \bar{\mathbf{H}} \bar{\mathbf{x}}_{\ell,s} \right\|_F^2 \right\} \quad (13)$$

where Θ_R is the subset of points (ℓ, s) for $\ell \in \{1, 2, \dots N_t\}$ and $s \in \{s_1, s_2, \dots s_M\}$ in the transmit search space that lie inside a sphere with radius R and centered around the received signal $\bar{\mathbf{y}}$. In formulas, the subset Θ_R can be written as shown in (14) on top of this page, where we have defined: i) $\bar{\mathbf{z}} = \mathbf{Q}_1^T \bar{\mathbf{y}}$, ii) $R_Q^2 = R^2 - \left\| \bar{\mathbf{Q}}_2^T \bar{\mathbf{y}} \right\|_F^2$, iii) $\bar{z}_i$ and $\bar{x}_i (\ell, s)$ are the i–th entry of vectors $\bar{\mathbf{z}}$ and $\bar{\mathbf{x}}_{\ell,s}$, respectively, and iv) $\bar{p}_{i,j}$ is the (i, j)–th entry of matrix $\bar{\mathbf{P}}$.

The key point behind the application of SD for reducing the computational complexity of ML–optimum detectors is the efficient computation of the subset Θ_R, which should avoid an exhaustive search in the whole transmit search space. As far as SM is concerned, these points can be computed in a very simple way, as summarized in *Lemma 1*.

Lemma 1: The subset of points Θ_R in (14) lie in the intervals:

$$\frac{-R_Q + \bar{z}_i}{\bar{p}_{i,i}} \leq \bar{x}_i (\ell, s) \leq \frac{R_Q + \bar{z}_i}{\bar{p}_{i,i}} \quad (15)$$

$$\frac{-R_Q + \bar{z}_{i,i+N_t}}{\bar{p}_{i,i}} \leq \bar{x}_i (\ell, s) \leq \frac{R_Q + \bar{z}_{i,i+N_t}}{\bar{p}_{i,i}} \quad (16)$$

for $i = 2N_t, 2N_t - 1, \dots, N_t + 1$ in (15), $i = N_t, N_t - 1, \dots, 1$ in (16), and we have defined $\bar{z}_{i,i+N_t} = \bar{z}_i - \bar{p}_{i,i+N_t} \bar{x}_{i+N_t} (\ell, s)$.

Proof: This result can be obtained as follows.

1) First, we note that a necessary condition that the points of the transmit search space need to satisfy to belong to the subset Θ_R in (14) is (for all $i = 1, 2, \dots, 2N_t$):

$$\left(\bar{z}_i - \sum_{j=i}^{2N_t} \bar{p}_{i,j} \bar{x}_j (\ell, s) \right)^2 \leq R_Q^2 \quad (17)$$

which is a condition similar to conventional SD algorithms [7].

2) Second, we need to take into account that in SM only a single antenna is active at any time instance. In the equivalent real–valued signal model in (10) and (13), this is equivalent to having only two, out of $2N_t$, non–zero entries in the signal vectors $\bar{\mathbf{x}}_{\ell_t, s_t}$ and $\bar{\mathbf{x}}_{\ell,s}$, respectively.

By taking this remark into account, it follows that: a) if $i = N_t + 1, N_t + 2, \dots, 2N_t$, then only the imaginary part of $\bar{\mathbf{x}}_{\ell,s}$ plays a role in (15), and, thus, only one entry $\bar{x}_j (\ell, s)$ can be non–zero; and b) if $i = 1, 2, \dots, N_t$, then both real and imaginary parts of $\bar{\mathbf{x}}_{\ell,s}$ play a role in (16), and, thus, only two entries $\bar{x}_j (\ell, s)$ can be non–zero. The considerations in a) and b) lead to the intervals in (15) and (16), respectively, which are directly obtained by solving the inequality in (17). $\qquad \square$

By comparing the intervals in (15) and (16) with those of a conventional SD [7], we notice that (15) and (16) are much simpler, and this is due to the fact that in SM there is only one active antenna element, while in conventional SDs all the antennas transmit simultaneously. Note that, as (16) depends implicitly on (15), this means that (16) needs to be computed for all the points that lie in the interval in (15).

With respect to the Rx–SD scheme, the Tx–SD scheme foresees some pre–computations to estimate the points that lie inside the sphere of radius R. These additional computations are carefully taken into account in the analysis of the computational complexity of the Tx–SD scheme and its comparison with the ML–optimum detector (see Section IV). Furthermore, we note that the radius R in (14) can still be computed from (9).

B. C–SD Detector

In Section II-D and Section III-A, we have seen that Rx–SD and Tx–SD aim at reducing the complexity of the ML–optimum detector by minimizing the size of the receive and transmit search spaces, respectively. So, it is natural to combine both of them to further decrease the receiver complexity by reducing the size of both the search spaces. The proposed C–SD is a two–step detector that works as follows:

1) First, the Tx–SD algorithm is used to reduce the transmit search space. The subset of points Θ_R is computed as shown in (14).
2) Second, the Rx–SD algorithm is used to reduce the receive search space. More specifically, Rx–SD is applied only on the subset of points Θ_R computed in the step above.

In formulas, we have:

$$\left[\hat{\ell}_t^{(\text{C-SD})}, \hat{s}_t^{(\text{C-SD})} \right] = \arg\min_{(\ell,s) \in \Theta_R} \left\{ \sum_{r=1}^{\tilde{N}_r(\ell,s)} |y_r - h_{\ell,r} s|^2 \right\} \quad (18)$$

where Θ_R and $\tilde{N}_r (\ell, s)$ are computed by using (14) and (7), respectively.

IV. ANALYSIS OF COMPUTATIONAL COMPLEXITY

In this section, we analyze the computational complexity of Tx–SD and C–SD algorithms by taking into account the pre–computations needed to estimate the points of the reduced search space. The complexity is here computed as the number of real multiplications needed by each algorithm. Furthermore, we note that when divisions are involved, we take them into account and assume they have the same complexity as multiplications.

A. Tx–SD

The computational complexity of Tx–SD can be estimated by taking into account that:

1) The QR factorization in (12), when computed by using the Householder algorithm [14], requires a number of real multiplications equal to:

$$\mathcal{C}_{QR} = \sum_{\xi=1}^{N} \left[2f\left(\xi\right) + f^2\left(\xi\right) + 2f^3\left(\xi\right) + 1 \right] - N_r^3 \quad (19)$$

where $f\left(\xi\right) = N_r + 1 - \xi$ and $N = \min\left\{N_r - 1, N_t\right\}$.

2) The computation of $\bar{\mathbf{z}} = \mathbf{Q}_1^T \bar{\mathbf{y}}$ in (14) needs $4N_t N_r$ real multiplications, i.e., $\mathcal{C}_{\bar{\mathbf{z}}} = 4N_t N_r$.

3) The computation of $\left\|\bar{\mathbf{Q}}_2^T \bar{\mathbf{y}}\right\|_{\mathrm{F}}^2$ in (14) needs $\mathcal{C}_{\left\|\bar{\mathbf{Q}}_2^T \bar{\mathbf{y}}\right\|_{\mathrm{F}}^2} = 2N_r\left(2N_r - 2N_t + 1\right)$ real multiplications, i.e., $\mathcal{C}_{\left\|\bar{\mathbf{Q}}_2^T \bar{\mathbf{y}}\right\|_{\mathrm{F}}^2} = 2N_r\left(2N_r - 2N_t + 1\right)$.

4) The computation of the intervals in (15) and (16) requires: i) $2N_t$ real divisions to compute (15) for N_t antenna indexes; and ii) 1 real multiplication and 2 real divisions for a number of times equal to the total number of points that satisfy (15). In the worst–case scenario, the number of points computed in (15) is equal to the size, M_I, of the imaginary constellation diagram composing the QAM constellation symbol. Accordingly, we have $M_I N_t$ real multiplications and $2M_I N_t$ real divisions, and, thus, the complexity of computing the intervals in (15) and (16) can be upper–bounded by $\mathcal{C}_{\mathrm{interval}} = 2N_t + 3M_I N_t$.

In summary, the analytical complexity resulting from the computation of the points in the subset Θ_R can be upper–bounded by:

$$\mathcal{C}_{\Theta_R} \leq \mathcal{C}_{QR} + \mathcal{C}_{\bar{\mathbf{z}}} + \mathcal{C}_{\left\|\bar{\mathbf{Q}}_2^T \bar{\mathbf{y}}\right\|_{\mathrm{F}}^2} + \mathcal{C}_{\mathrm{interval}} \quad (20)$$

Since (13) requires 8 real multiplications for each computed Euclidean distance, it follows that the computational complexity of the Tx–SD receiver can be upper–bounded as follows:

$$\mathcal{C}_{\mathrm{Tx-SD}} \leq \mathcal{C}_{\Theta_R} + 8N_r \mathrm{card}\left\{\Theta_R\right\} \quad (21)$$

where $\mathrm{card}\left\{\cdot\right\}$ denotes the cardinality, i.e., the number of points, in a set.

B. C–SD

The computational complexity of C–SD follows immediately from (8) and (21), as follows:

$$\mathcal{C}_{\mathrm{C-SD}} \leq \mathcal{C}_{\Theta_R} + 8 \sum_{(\ell,s) \in \Theta_R} \tilde{N}_r\left(\ell, s\right) \quad (22)$$

V. NUMERICAL RESULTS

In this section, Monte Carlo simulation results for at least 10^6 channel realisation are shown to compare the performance and computational complexity of ML–optimum and SD–based receivers. The numerical examples are obtained by assuming the system model in Section II. Furthermore, the radius R is chosen as described in (9) with $\varepsilon = 10^{-6}$.

In Fig. 1, the BEP averaged over Rayleigh fading is shown by considering two different constellation sizes, and $N_t = N_r = 4$. We notice that all the SDs have the same performance, and all of them overlap with the ML–optimum detector. As expected, the performance of SM improves when M decreases.

In Fig. 2, Fig. 3, and Fig. 4, the computational complexity $\mathcal{C}$ is compared according to the estimates provided in the sections above. In particular, the figures show the relative computational complexity of the SDs with respect to the ML–optimum detector, i.e., $C_{\mathrm{rel}}\left(\%\right) = 100 \times \left(C_{\mathrm{ML}} - C_{\mathrm{SD}}\right)/C_{\mathrm{ML}}$. In Fig. 2, we observe that C–SD is always better than Tx–SD, while it is better than Rx–SD only for high SNRs. The reason for this latter result is due to the additional pre–computations required by both C–SD and Tx–SD solutions. As mentioned in Section I, we notice that the Rx–SD scheme is more effective in reducing the complexity when the number of antennas at the receiver is large.

Figures 3 and 4 show an interesting setups where it is clearly highlighted that none of the proposed SDs is *a priori* superior to the others, and that the best detector to use actually depends on both the MIMO setup and the SNR. In particular, on the right–hand side of Fig. 3, we notice that the best receiver to use is the C–SD, while if we look at the left–hand side of Fig. 3 and Fig. 4 we notice that the best receiver to use is always the RX–SD. The reason is that in the former case the transmit search space is large enough to compensate for the pre–computations required by Tx–SD and C–SD receivers. On the contrary, in the latter case the transmit search space is not very large, and, thus, there is not too much gain in reducing it.

Finally, the right–hand side of Fig. 4 shows that Rx–SD and C–SD are almost equivalent, as both a large transmit search space and a large number of antennas at the receiver are considered.

In summary, we can conclude that Rx–SD turns out to be the best choice when the number of antennas at the receiver is large, as it can significantly reduce the receiver search space without any pre–computations. On the other hand, C–SD turns out to be the best choice when the transmit search space is very large and the number of antennas at the receiver is low. It is interesting to note that the Tx–SD scheme is never the best choice in the analyzed system setup. Since this latter decoder is the most similar to conventional SDs for MIMO systems, a main contribution of this paper is to have highlighted the importance of reducing both transmit and receive search spaces to get a significant reduction in receiver complexity with performance guarantees.
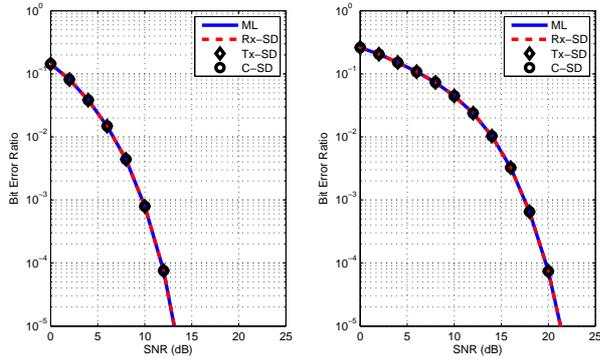
Fig. 1. BEP against the SNR. $N_t = N_r = 4$, (left) $M = 8$; (right) $M = 64$.

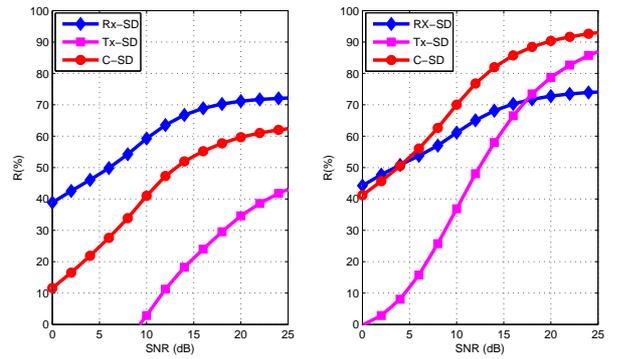

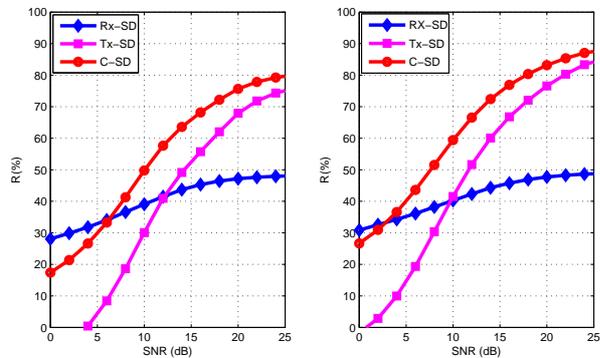Fig. 3. Computational complexity against the SNR. $N_t = N_r = 4$, (left) $M = 8$; (right) $M = 64$.

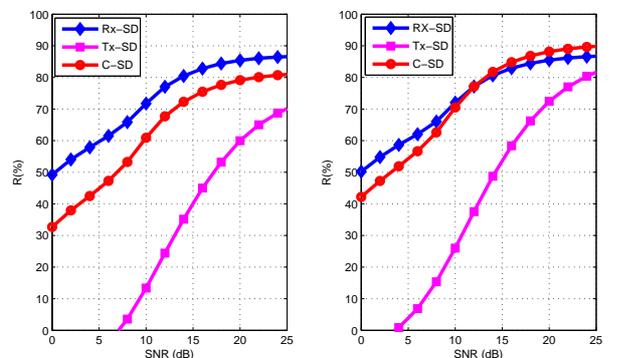Fig. 2. Computational complexity against the SNR. $N_t = N_r = 2$, (left) $M = 8$; (right) $M = 16$.

Fig. 4. Computational complexity against the SNR. $N_t = N_r = 8$, (left) $M = 32$; (right) $M = 64$.

VI. CONCLUSION

In this paper, we have introduced and analyzed the performance/complexity trade–off of three SDs for SM. We have shown that the proposed solutions provide a substantial reduction of the computational complexity while retaining the same BEP as the ML–optimum detector. Numerical results have highlighted that no SD is superior to the others, and that the best solution to use depends on the MIMO setup, *i.e.*, the triple (N_t, N_r, M), and the SNR at the receiver. In general, Rx–SD is the best choice for large N_r, and C–SD is the best option when either N_r is low or M is large. Overall, analysis and results shown in this paper confirm that SD can be a viable and effective solution to reduce the complexity of SM without deteriorating its BER performance.

REFERENCES

[1] R. Mesleh, H. Haas, S. Sinanović, C. W. Ahn, and S. Yun, "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228 – 2241, Jul. 2008.

[2] J. Jeganathan, A. Ghrayeb, and L. Szczecinski, "Spatial modulation: optimal detection and performance analysis," *IEEE Commun. Lett.*, vol. 12, no. 8, pp. 545–547, Aug. 2008.

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

[4] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Union Radio-Scientifique Internationale (URSI) Intern. Symp. on Signals, Systems, and Electronics (ISSSE)*, Sep. 29– Oct. 2, 1998, pp. 295–300.

[5] M. D. Renzo and H. Haas, "Performance analysis of spatial modulation," in *IEEE International Conference on Communication and Networking in China (CHINACOM)*, Beijing, China, Aug. 2010, pp. 1–7, (invited paper).

[6] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Commun. Lett.*, vol. 12, no. 4, pp. 225 –227, Apr. 2008.

[7] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. expected complexity," *IEEE Trans. on Signal Process.*, vol. 53, no. 8, pp. 2806– 2818, Aug. 2005.

[8] A. Younis, R. Mesleh, H. Haas, and P. Grant, "Reduced complexity sphere decoder for spatial modulation detection receivers," in *IEEE Globecom 2010 - Wireless Communications Symposium (GC10 - WC)*, Miami, Florida, USA, Dec. 2010.

[9] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1639–1642, Jul. 1999.

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

[11] P. Wang and T. Le-Ngoc, "A low-complexity generalized sphere decoding approach for underdetermined linear communication systems: performance and complexity evaluation," *IEEE Trans. on Commun.*, vol. 57, no. 11, pp. 3376 –3388, Nov 2009.

[12] ——, "A list sphere decoding algorithm with improved radius setting strategies," *Wireless Personal Communications*, pp. 1–12, 2010, 10.1007/s11277-010-0018-4. [Online]. Available: http://dx.doi.org/10. 1007/s11277-010-0018-4

[13] H. V. T. Kailath and B. Hassibi, *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge University Press, 2006, ch. MIMO Receive Algorithms.

[14] S. Aubert, F. Nouvel, and A. Nafkha, "Complexity gain of QR decomposition based sphere decoder in LTE receivers," in *IEEE 70th Vehicular Technology Conference Fall (VTC 2009-Fall)*, Sep. 2009, pp. 1–5.