Reduced Complexity Sphere Decoding via a Reordered Lattice Representation

Luay Azzam, *Student Member, IEEE,* and Ender Ayanoglu, *Fellow, IEEE*

***Abstract*—In this letter, we propose a reordering of the channel representation for Sphere Decoding (SD) where the real and imaginary parts of each jointly detected symbol are decoded independently. Making use of the proposed structure along with a scalar quantization technique, we reduce the decoding complexity substantially. We show that this approach achieves 85% reduction in the overall complexity compared to the conventional SD for a 2×2 system, and 92% reduction for the 4×4 and 6×6 cases at low SNR values, and almost 50% at high SNR, thus relaxing the requirements for hardware implementation.**

***Index Terms*—Maximum-likelihood detection, multiple-input multiple-output channels, sphere decoding.**

I. Introduction

CONSIDER a MIMO system with N transmit and M receive antennas. The received signal is given by

$$y = Hs + v \tag{1}$$

where $y \in \mathbb{C}^M$ is an M-dimensional received complex vector, $H \in \mathbb{C}^{M \times N}$ is the channel matrix, $s \in \mathbb{C}^N$ is an N-dimensional transmitted complex vector whose entries have real and imaginary parts that are integers, $v \in \mathbb{C}^M$ is the i.i.d. complex additive white Gaussian noise (AWGN) vector with zero-mean and covariance matrix $\sigma^2 I$. Usually, the elements of the vector s are constrained to a finite set Ω where $\Omega \subset \mathbb{Z}^{2N}$, e.g., $\Omega = \{-3, -1, 1, 3\}^{2N}$ for 16-QAM (quadrature amplitude modulation) where $\mathbb{Z}$ and $\mathbb{C}$ denote the sets of integers and complex numbers, respectively.

Assuming that the receiver has perfect knowledge of the channel H, different algorithms have been implemented to separate the data streams corresponding to N transmit antennas [1], [2], [3]. Among these algorithms, Maximum Likelihood Decoding (MLD) is the optimum one. However, in MIMO systems, the MLD problem becomes more complicated as the constellation size is larger [4]. Sphere decoding [5], on the other hand, or the Viterbo-Boutros algorithm [6], reduces the computational complexity for the class of computationally hard combinatorial problems that arise in MLD [7], [8].

In this letter we improve the SD complexity efficiency by reducing the number of arithmetic operations (mainly the number of multiplications) required by the SD algorithm. The reduction of the number of arithmetic operations is accomplished by introducing a simple and proper lattice representation, and adopting a technique similar to the well-known K-best algorithm [9] with some modifications, as well as incorporating rounding at a certain level of the SD search which significantly reduces the number of nodes to be visited throughout the search process. The validity of this lattice representation is proved for any matrix size by using the QR decomposition and the Gram-Schmidt orthogonalization procedure which distinguishes it from similar work in the literature [10], [11]. In addition, rounding to the nearest neighbor or quantization is employed that maintains near optimal performance while significantly reducing the complexity even when used for spatial multiplexing systems having four or more transmit antennas. This results in a simpler system that does not compromise performance, unlike similar previous work in the literature. Finally, it is important to mention that searching the lattice points using this formulation can be performed in parallel, since the proposed structure enables decoding the real and imaginary parts of each symbol independently and at the same time.

II. Problem Definition and the Conventional Sphere Decoder

Assuming H is known at the receiver, the MLD solution of (1) is given by

$$\hat{s} = \arg \min_{s \in \Omega^N} ||y - Hs||^2. \tag{2}$$

Solving (2) becomes complicated for large constellations and large N. Therefore, instead of searching the whole space defined by all combinations drawn by the set Ω, SD solves this problem by searching only over those lattice points or combinations that lie inside a sphere centered around the received vector y and of radius d [5], [12], [13]. Introducing this constraint on (2) changes the problem to

$$\hat{s} = \arg \min_{s \in \Omega^N} ||y - Hs||^2 < d^2. \tag{3}$$

A frequently used solution for the QAM-modulated complex signal model given in (3) is to decompose the N-dimensional problem into a $2N$-dimensional real-valued problem, which then can be written as

$$\begin{bmatrix} \Re\{y\} \\ \Im\{y\} \end{bmatrix} = \begin{bmatrix} \Re\{H\} & -\Im\{H\} \\ \Im\{H\} & \Re\{H\} \end{bmatrix} \begin{bmatrix} \Re\{s\} \\ \Im\{s\} \end{bmatrix} + \begin{bmatrix} \Re\{v\} \\ \Im\{v\} \end{bmatrix} \tag{4}$$

where $\Re\{y\}$ and $\Im\{y\}$ denote the real and imaginary parts of y, respectively. Assuming $N = M$ in the sequel, and introducing the QR decomposition of H, where R is an upper triangular matrix, and the matrix Q is unitary, (3) can be written as

$$\hat{s} = \arg \min_{s \in \Omega^{2N}} ||\bar{y} - Rs||^2 < d^2 \tag{5}$$

Paper approved by C. Tellambura, the Editor for Multicarrier Systems of the IEEE Communications Society. Manuscript received May 23, 2007; revised November 19, 2007 and April 15, 2008.

The authors are with the Center for Pervasive Communications and Computing, Department of Electrical Engineering and Computer Science, University of California, Irvine, Irvine, CA 92697-2625 (e-mail: luay125@hotmail.com, lazzam@uci.edu).

Digital Object Identifier 10.1109/TCOMM.2009.09.070238

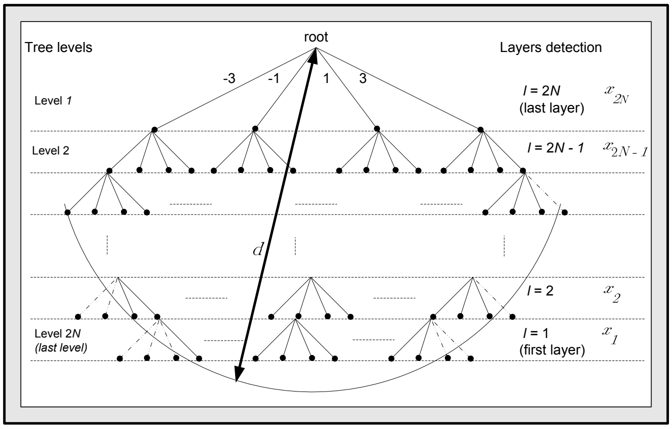

Fig. 1. Tree search example for a 16-QAM showing sphere radius, tree levels and detection layers.

where $\bar{y} = Q^H y$. Let R=$[r_{i,j}]_{2N\text{x}2N}$ and note that R is upper triangular. Now to solve (5), SD algorithm constructs a tree as shown in Figure 1, where the branches coming out of each node correspond to the elements drawn by the set Ω. It then executes the decoding process starting from the last layer ($l = 2N$) which matches the first level in the tree, calculating the partial weight $|\Im\{\bar{y}_N\} - r_{2N,2N}\Im\{s_N\}|^2$, and working its way down the tree in a similar way to the successive interference cancelation technique, until decoding the first layer by calculating the corresponding partial weight $|\Re\{\bar{y}_1\} - r_{1,1}\Re\{s_1\} + ... + \Re\{\bar{y}_N\} - r_{1,N}\Re\{s_N\} + \Im\{\bar{y}_1\} - r_{1,N+1}\Im\{s_1\} + ... + \Im\{\bar{y}_N\} - r_{1,2N}\Im\{s_N\}|^2$.

The weight of any node in the tree is the summation of the partial weight of that node and the weight of its parent node. If that weight exceeds the square of the sphere radius d^2, the algorithm prunes the corresponding branch, declaring it as an improbable way to a candidate solution. In other words, all branches that lead to a solution that is outside the sphere are pruned at some level of the tree. These branches are shown as dashed lines in Figure 1. Whenever a valid lattice point at the last level of the tree is found within the sphere, the square of the sphere radius d^2 is set to the newly found point weight, thus reducing the search space for finding other candidate solutions. Finally, the leaf with the lowest weight is the survivor one, and the path along the tree from the root to that leaf represents the estimated solution.

To this end, it is important to emphasize the fact that the complexity of this algorithm, although much lower than MLD, is still exponential at low SNR. We aim to reduce that complexity as will be shown subsequently.

III. Proposed Algorithm

The lattice representation given in (4) imposes a major restriction on the tree search algorithm. Specifically, the search has to be executed serially from one level to another on the tree. This can be made clearer by writing the partial metric weight formula as

$$w_l(x^{(l)}) = w_{l+1}(x^{(l+1)}) + |\hat{y}_l - \sum_{k=l}^{2N} r_{l,k} x_k|^2 \qquad (6)$$

with $l = 2N, 2N-1, \ldots, 1$, $w_{2N+1}(x^{(2N+1)}) = 0$ and where $\{x_1, x_2, ..., x_N\}$, $\{x_{N+1}, x_{N+2}, ..., x_{2N}\}$ are the real and imaginary parts of $\{s_1, s_2, ..., s_N\}$ respectively, see Figure 1. According to this representation, it is impossible, for instance, to calculate $\sum_{k=l}^{2N} r_{l,k} x_k$ in (6) for a node that lies at level ($l = 2N - 1$) without assigning an estimate for x_{2N}. This approach results in two related drawbacks. First, the decoding of any x_l requires an estimate value for all preceding x_j for $j = l+1, ..., 2N$. Secondly, there is no room for parallel computations since the structure of the tree search is sequential.

The main contribution in this letter is that we relax the tree search structure making it more flexible for parallelism, and at the same time reducing the number of computations required at each node by making the decoding of every two adjacent levels in the tree totally independent of each other. This is achieved using the important observation that results from applying the QR decomposition to the proposed lattice representation. We further reduce the complexity by keeping the best K symbols which have lowest weights in a way similar to the well-known K-best technique, but with a number of differences that are to be discussed subsequently. We propose the new algorithm in detail in the following subsections.

A. Proposed Lattice Representation

We start by reshaping the channel matrix representation given in (4) in the following form

$$\tilde{H} = \begin{bmatrix} \Re(H_{1,1}) & -\Im(H_{1,1}) & \cdots & \Re(H_{1,N}) & -\Im(H_{1,N}) \\ \Im(H_{1,1}) & \Re(H_{1,1}) & \cdots & \Im(H_{1,N}) & \Re(H_{1,N}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Re(H_{N,1}) & -\Im(H_{N,1}) & \cdots & \Re(H_{N,N}) & -\Im(H_{N,N}) \\ \Im(H_{N,1}) & \Re(H_{N,1}) & \cdots & \Im(H_{N,N}) & \Re(H_{N,N}) \end{bmatrix} \qquad (7)$$

where $H_{m,n}$ is the i.i.d. complex path gain from transmit antenna n to receive antenna m. By careful observation of the columns of $\tilde{H}$ starting from the left hand side, and defining each pair of columns as one set, we note that the columns in each set are orthogonal, a property that has a substantial effect on the structure of the problem. Using this channel representation changes the order of detection of the transmitted symbols from

$$\hat{s} = \begin{bmatrix} \Re(s_1) & \cdots & \Re(s_N) & \Im(s_1) & \cdots & \Im(s_N) \end{bmatrix}^T \qquad (8)$$

to the following order

$$\hat{s} = \begin{bmatrix} \Re(s_1) & \Im(s_1) & \cdots & \Re(s_N) & \Im(s_N) \end{bmatrix}^T. \qquad (9)$$

This means that the first and second levels of the search tree (see Figure (1)) correspond to the real and imaginary parts of s_N, unlike conventional SD, where these levels correspond to the imaginary parts of s_N and s_{N-1}, respectively. This structure becomes advantageous after applying the QR decomposition to $\tilde{H}$. This is formalized in the following theorem.

Theorem 1: Applying QR decomposition to the channel matrix $\tilde{H}$ which has the aforementioned orthogonal properties among its columns produces an upper triangular matrix R whose elements $r_{k,k+1}$ are all zero for $k = 1, 3, ..., 2N-1$.

Proof: Let

$$\tilde{H} = \begin{bmatrix} \tilde{\mathbf{h}}_1 & \tilde{\mathbf{h}}_2 & \cdots & \tilde{\mathbf{h}}_{2N} \end{bmatrix} \tag{10}$$

where $\tilde{\mathbf{h}}_k$ is the kth column of $\tilde{H}$. Recalling the Gram-Schmidt algorithm, we define $\mathbf{u}_1 = \tilde{\mathbf{h}}_1$, and then $\mathbf{u}_k = \tilde{\mathbf{h}}_k - \sum_{j=1}^{k-1} \phi_{\mathbf{u}_j} \tilde{\mathbf{h}}_k$ for $k = 2, 3, ..., 2N$, where $\phi_{\mathbf{u}_j} \tilde{\mathbf{h}}_k$ is the projection of $\tilde{\mathbf{h}}_k$ onto $\mathbf{u}_j$ defined by

$$\phi_{\mathbf{u}_j} \tilde{\mathbf{h}}_k = \frac{\langle \tilde{\mathbf{h}}_k, \tilde{\mathbf{u}}_j \rangle}{\langle \tilde{\mathbf{u}}_j, \tilde{\mathbf{u}}_j \rangle} \tilde{\mathbf{u}}_j. \tag{11}$$

We also define $\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$ for $k = 1, 2, \ldots, 2N$, and rewrite the column vectors of the channel matrix $\tilde{H}$ in the equations form as

$\tilde{\mathbf{h}}_1 = \mathbf{e}_1 \|\mathbf{u}_1\|$

$\tilde{\mathbf{h}}_2 = \phi_{\mathbf{u}_1} \tilde{\mathbf{h}}_2 + \mathbf{e}_2 \|\mathbf{u}_2\|$

$\vdots$

$\tilde{\mathbf{h}}_k = \sum_{j=1}^{k-1} \phi_{\mathbf{u}_j} \tilde{\mathbf{h}}_k + \mathbf{e}_k \|\mathbf{u}_k\|$.

Now, writing $Q = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \cdots & \mathbf{e}_n \end{bmatrix}$, we have these equations in the matrix form as

$$Q \begin{bmatrix} \|\mathbf{u}_1\| & \langle \mathbf{e}_1, \tilde{\mathbf{h}}_2 \rangle & \cdots & \langle \mathbf{e}_1, \tilde{\mathbf{h}}_n \rangle \\ 0 & \|\mathbf{u}_2\| & \cdots & \langle \mathbf{e}_2, \tilde{\mathbf{h}}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \|\mathbf{u}_n\| \end{bmatrix}. \tag{12}$$

Obviously, the matrix Q is unitary, and the matrix on the right is the upper triangular R matrix of the QR decomposition. Now our task is to show that the terms $\langle \mathbf{e}_k, \tilde{\mathbf{h}}_{k+1} \rangle$ are zero for $k = 1, 3, \ldots, 2N-1$. Three observations conclude the proof.

First, since $\tilde{\mathbf{h}}_k$ and $\tilde{\mathbf{h}}_{k+1}$ are orthogonal for $k = 1, 3, \ldots, 2N-1$, then $\phi_{\mathbf{u}_k} \tilde{\mathbf{h}}_{k+1} = \phi_{\mathbf{u}_{k+1}} \tilde{\mathbf{h}}_k = 0$ for same k.

Second, the projection of $\mathbf{u}_m$ for $m = 1, 3, \ldots, k-2$ on the columns $\tilde{\mathbf{h}}_k$ and $\tilde{\mathbf{h}}_{k+1}$ is equal to the projection of $\mathbf{u}_{m+1}$ on the columns $\tilde{\mathbf{h}}_{k+1}$ and -$\tilde{\mathbf{h}}_k$, respectively. To formalize this

$$\langle \mathbf{u}_m, \tilde{\mathbf{h}}_k \rangle = \langle \mathbf{u}_{m+1}, \tilde{\mathbf{h}}_{k+1} \rangle$$

and,

$$\langle \mathbf{u}_m, \tilde{\mathbf{h}}_{k+1} \rangle = -\langle \mathbf{u}_{m+1}, \tilde{\mathbf{h}}_k \rangle$$

for $k = 1, 3, \ldots, 2N-1$ and $m = 1, 3, \ldots, k-2$. This property becomes obvious by using the first observation and revisiting the special structure of (7).

Third, making use of the first two observations, and noting that $||\tilde{\mathbf{h}}_k||$=$||\tilde{\mathbf{h}}_{k+1}||$ for $k = 1, 3, \ldots, 2N-1$, it can be easily shown that $||\mathbf{u}_k||$=$||\mathbf{u}_{k+1}||$ for same k. Then,

$$\begin{aligned} \langle \mathbf{e}_k, \tilde{\mathbf{h}}_{k+1} \rangle =& \langle \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}, \tilde{\mathbf{h}}_{k+1} \rangle \\ =& \frac{1}{\|\mathbf{u}_k\|} \langle \tilde{\mathbf{h}}_k - \sum_{j=1}^{k-1} \phi_{\mathbf{u}_j} \tilde{\mathbf{h}}_k, \tilde{\mathbf{h}}_{k+1} \rangle \\ =& \frac{1}{\|\mathbf{u}_k\|} (\langle \tilde{\mathbf{h}}_k, \tilde{\mathbf{h}}_{k+1} \rangle - \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_1 \rangle \langle \mathbf{u}_1, \tilde{\mathbf{h}}_{k+1} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} - \\ & \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_2 \rangle \langle \mathbf{u}_2, \tilde{\mathbf{h}}_{k+1} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} - \ldots - \\ & \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_{k-2} \rangle \langle \mathbf{u}_{k-2}, \tilde{\mathbf{h}}_{k+1} \rangle}{\langle \mathbf{u}_{k-1}, \mathbf{u}_{k-1} \rangle} - \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_{k-1} \rangle \langle \mathbf{u}_{k-1}, \tilde{\mathbf{h}}_{k+1} \rangle}{\langle \mathbf{u}_{k-2}, \mathbf{u}_{k-2} \rangle}) \end{aligned} \tag{13}$$

Now, applying the above observations to (13), we get

$$\begin{aligned} \langle \mathbf{e}_k, \tilde{\mathbf{h}}_{k+1} \rangle =& \frac{1}{\|\mathbf{u}_k\|} (0 - \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_1 \rangle \langle \mathbf{u}_1, \tilde{\mathbf{h}}_{k+1} \rangle}{\|\mathbf{u}_1\|^2} - \\ & \frac{-\langle \mathbf{u}_1, \tilde{\mathbf{h}}_{k+1} \rangle \langle \tilde{\mathbf{h}}_k, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} - \ldots - \\ & \frac{\langle \tilde{\mathbf{h}}_k, \mathbf{u}_{k-2} \rangle \langle \mathbf{u}_{k-2}, \tilde{\mathbf{h}}_{k+1} \rangle}{\|\mathbf{u}_{k-1}\|^2} - \frac{-\langle \mathbf{u}_{k-2}, \tilde{\mathbf{h}}_{k+1} \rangle \langle \tilde{\mathbf{h}}_k, \mathbf{u}_{k-2} \rangle}{\|\mathbf{u}_{k-1}\|^2}) \\ =& 0. \end{aligned}$$

This concludes the proof. ■

The locations of these zeros are very important since they introduce orthogonality between the real and imaginary parts of every detected symbol. In this context, the SD algorithm executes in the following way. The partial metric weight $|\hat{y}_{2N} - r_{2N,2N}\hat{x}_{2N}|^2$ for the μ nodes in the first level of the tree is computed, where μ is the number of elements in Ω (μ equals 4 for the 16-QAM example shown in Figure 1). This metric is then checked against the specified sphere radius d^2. If the weight of any node is greater than the sphere radius, then the corresponding branch is pruned. Otherwise, the metric value is saved for the next step. At the same time, another set of μ partial metric computations of the form $|\hat{y}_{2N-1} - r_{2N-1,2N-1}\hat{x}_{2N-1}|^2$ takes place at the second level, since these two levels are independent as stated above. These metrics are checked against d^2 in a similar way to that carried out in the first level. The weights of the survivor nodes from both levels are summed up and the summation is checked against the sphere constraint, ending up with a set of survivor symbols $\hat{s}_N$. The estimation of the remaining $N-1$ symbols is performed recursively in a similar way executing two levels of the tree at a time.

To further reduce the complexity it is often useful to apply some bounding techniques on the enumeration of the nodes to be considered at the tree levels. This is done by recursively defining upper and lower bounds on $\hat{x}_i$ for $i = 2N, 2N-1, \ldots, 1$ where *at most* μ elements for each $\hat{x}_i$ belong to the interval defined by these bounds. This efficient technique has been considered before in [8]. For a fair comparison, we either incorporate this technique in conventional SD and our proposed algorithm or dismiss it for both.

Finally, it is important to notice that the performance of the proposed algorithm using the proposed lattice representation is exactly the same as conventional SD, whereas the complexity is reduced significantly as will be shown.

B. Rounding to the Nearest Neighbor (Quantization)

The estimation of the symbols can be carried out recursively by rounding (or quantizing) to the nearest constellation element in Ω. This heuristic approach together with the term quantization was used before for decoding layered space-time codes (LST) [14]. However, applying rounding for detecting all or most of the transmitted symbols would cause high performance loss especially when N is large ($N \geq 4$). Note that in the execution of the detection algorithm with rounding, variables s_2, s_3, and s_4 (for $N = 4$) are determined using MLD metric and only s_1 is detected by rounding. As a result, the performance degradation due to quantization with respect to MLD is very small. On the other hand, due to the expanding structure of SD search, computational complexity gains are substantial. Therefore, a careful combination of the proposed algorithm in the previous subsection, and the rounding technique proposed here will give substantially improved results in terms of performance and complexity. The use of rounding technique has a large impact on the reduction of the number of arithmetic operations required at the receiver.

In order to make up for the performance loss caused by rounding, we introduce an adaptive K-best technique in the middle levels of the tree. This results in near optimal performance (< 1 dB loss) and keeps the decoding complexity substantially below that required for conventional SD.

C. Adaptive K-BEST

The K-best algorithm (equivalent to the M-algorithm [15]) is a breadth-first search approach that has been widely used in the VLSI implementation of lattice decoders [16]-[21]. In K-best we only expand K nodes which have the smallest accumulated partial weights at each level in the tree search. Here we propose similar approach but with the following differences

- The value of K becomes smaller as we further traverse down the tree.
- The best K criterion is invoked every two levels in the tree search unlike conventional K-best where it is invoked every level. This gives better performance results.

Finally, it is worth mentioning that the value of K is chosen according to the modulation scheme and the number of antennas used. Suggested values are given in the next section and the corresponding simulation results are provided.

IV. Simulation Results

We consider 2×2, 4×4, and 6×6 cases using 16-QAM and 64-QAM modulation schemes. In all simulations we use ($d^2 = 2\sigma^2 N$ [22]) as the initial radius value (the radius choice problem has been widely treated in the literature, e.g., [6], [22], [23] and the references therein). Adaptive K-best is invoked only for $N = M = 4$ and $N = M = 6$. The proposed lattice representation is considered and quantization is applied at the very low levels of the tree. For instance, for $N = 6$ the tree has 12 levels (see Figure 1) and rounding is applied only at the last four levels. Moreover, on the tree

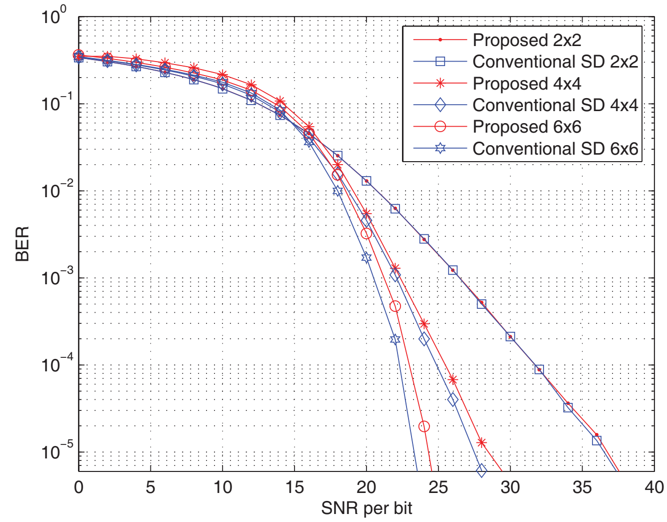

Fig. 2. BER vs SNR per bit for the proposed and conventional SD over a 2×2, 4×4, and 6×6 MIMO flat fading channel using 16-QAM.

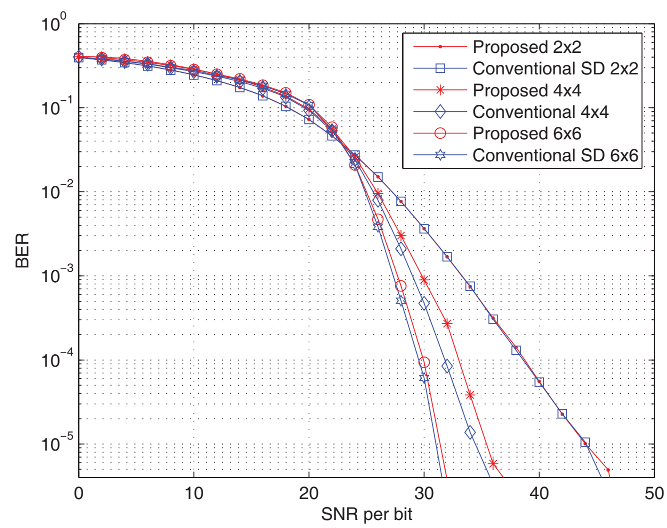

Fig. 3. BER vs SNR per bit for the proposed and conventional SD over a 2×2, 4×4, and 6×6 MIMO flat fading channel using 64-QAM.

levels that correspond to the detection of $\hat{s}_6$, $\hat{s}_5$, and $\hat{s}_4$, we invoke adaptive K-best at every two levels with values $16, 8$, and 4 for 16-QAM and $32, 32$, and 16 for 64-QAM.

Figure 2 and Figure 3 show the performance of the proposed algorithm versus conventional SD for 2×2, 4×4, and 6×6 systems using 16-QAM and 64-QAM modulation schemes, respectively. We observe that both algorithms have exactly the same performance for 2×2 and < 1 dB performance loss in the proposed for 4×4 and 6×6. This loss is due to the adoption of the K-best algorithm and the rounding process used.

A complexity comparison is given in Figure 4 and Figure 5. The complexity is measured in terms of the number of real multiplications required to jointly decode N transmitted symbols. Similar curves with same percentage of reductions in the complexity are obtained when we consider the number of real additions. Compared to conventional SD, the proposed algorithm reduces the complexity by 85% for a 2×2 system and 92% (95%) for 4×4 (6×6) at low SNR, and 45% (50%)

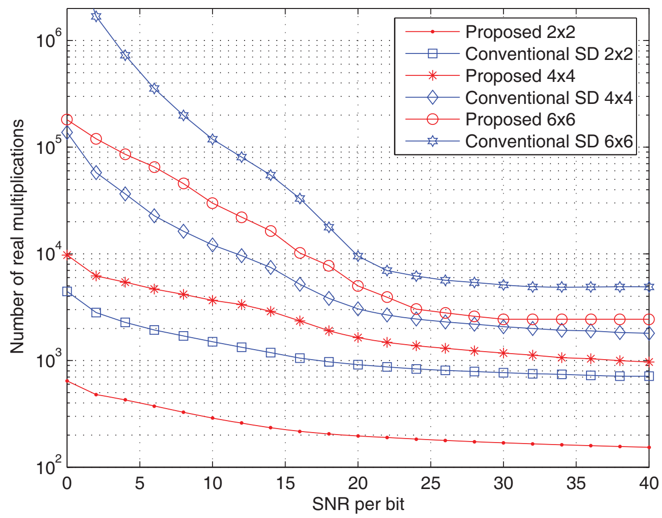

Fig. 4. Number of real multiplications vs SNR per bit for the proposed and conventional SD over a 2×2, 4×4, and 6×6 MIMO flat fading channel using 16-QAM.

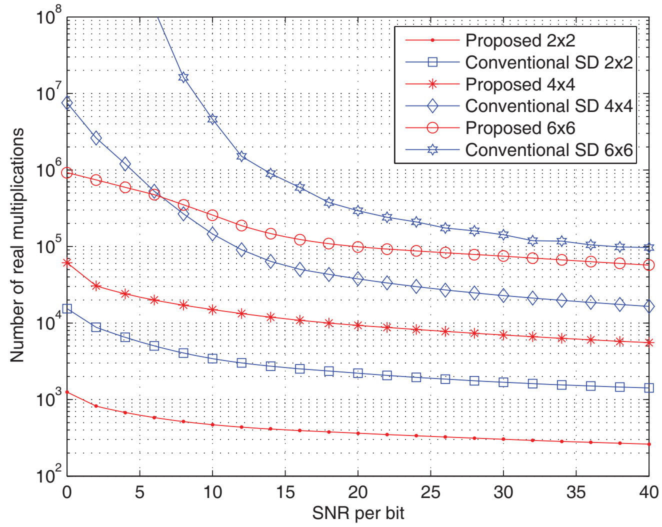

Fig. 5. Number of real multiplications vs SNR per bit for the proposed and conventional SD over a 2×2, 4×4, and 6×6 MIMO flat fading channel using 64-QAM.

for the same case at high SNR values.

The proposed lattice representation is a simple channel re-ordering technique that leads to the aforementioned complexity reduction. However, there are different ordering techniques that were proposed in the literature [3], [12] and yield to a reduced complexity SD as well. Therefore, we provide a complexity comparison between the conventional SD, SD with V-BLAST ZF-DFE ordering, SD with V-BLAST MMSE-DFE ordering, and our proposed algorithm for a 4×4 system employing 64-QAM . For a fair comparison, we determine the admissible sets at each level of the tree by applying the interval boundary conditions and considering only those points that belong to these intervals, as explained in [8]. As a result, we exclude the adoption of the adaptive K-best technique in our proposed algorithm, and apply quantization (rounding) in the last four levels of the tree.

The proposed algorithm provides a complexity gain of

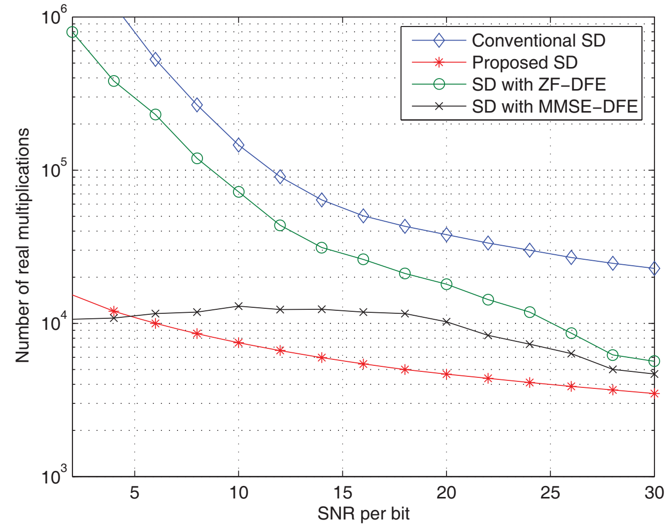

Fig. 6. Number of real multiplications vs SNR per bit for the conventional SD, proposed SD, SD with ZF-DFE ordering, and SD with MMSE-DFE ordering over a 4×4 MIMO flat fading channel using 64-QAM.

$> 90\%$ at low SNR and $> 35\%$ at high SNR values compared to SD with ZF-DFE ordering as shown in Figure 6. Considering SD with MMSE-DFE ordering, we see that for SNR values less than 5 dB, the proposed algorithm requires a slightly higher complexity. This complexity loss is reversed and becomes a gain reaching 50% at the intermediate values of SNR. In Figure 6, we only consider the complexity of the searching phase. The cost of the preprocessing phase depends only on the channel H. If the channel is constant for a long time, then the complexity of the preprocessing is insignificant compared to the overall complexity. However, if the channel changes arbitrarily then the complexity of the preprocessing may have a significant impact on the overall complexity [12].

To this end, it is to be noted that SD with ZF-DFE ordering provides the optimal performance as the conventional SD, whereas both SD with MMSE-DFE ordering and our proposed algorithm have an insignificant performance degradation.

A. Sphere Decoder versus Schnorr-Euchner Decoder

In the preceding discussion, we referred to the Viterbo-Boutros algorithm by SD. The Schnorr-Euchner (SE) algorithm [22], [24], on the other hand, has the same principle as SD, which is searching for the closest point inside a defined sphere. However, they differ in the way they find this point. In [22], it was shown that the complexity of SD and that of SE are very close to each other with a little advance for SE when the number of antennas is small. For a large number of antennas, this situation is reversed and SD becomes faster. In this work, we give a complexity comparison between both enumerations using our proposed algorithm.

Following [12], the comparison is carried out considering only the searching phase, noting that the complexity of the preprocessing phase for SE is higher than that for SD [22]. In Figure 7, we show the number of real multiplications required to decode the transmitted symbols as a function of the number of antennas at 20 dB for an uncoded system employing 16 QAM. Obviously, the complexity curves are very close to each

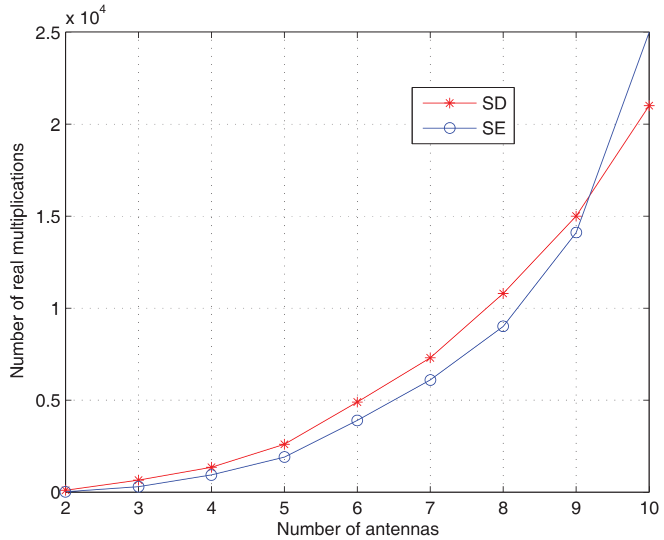

Fig. 7. Number of real multiplications vs number of antennas required for SE and SD at 20 dB considering an uncoded system with 16-QAM.

other for a number of antennas less than 9 with SE being a little faster. For more than 9 antennas, SD becomes much faster than SE. These results coincide with those obtained in [22]. As a result, we conclude that using either of these enumerations for conventional and proposed algorithms produces almost the same complexity gains explained above.

Finally, it is important to emphasize the fact that the proposed algorithm works for all square QAM modulation schemes and achieves similar results as those obtained for 16-QAM and 64-QAM.

V. Conclusions

In this letter, a general lattice representation via a simple channel ordering for sphere decoding was proposed. Rounding and adaptive K-best techniques are applied to enhance the proposed structure. For 2×2 systems, the performance of the proposed structure is the same as that for conventional SD, while it has < 1 dB loss for 4×4 and 6×6 cases. A complexity reduction of 80% is achieved for the 2×2 case, and more than 50% for the 4×4 and 6×6 cases.

Acknowledgment

We thank the reviewers for their valuable suggestions which improved the quality of this letter.

References

[1] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *IEEE J. Solid-State Circuits*, vol. 40, pp. 1566–1577, July 2005.

[2] J. Adeane, M. Rodrigues, I. Berenguer, and I. Wassell, "Improved detection methods for MIMO-OFDM-CDM communication systems," in *Proc. IEEE VTC*, vol. 3, pp. 1604–1608, Sept. 2004.

[3] G. Golden, G. Foschini, R. Valenzuela, and P. Wolniansky, "Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electron. Lett.*, vol. 35, no. 1, pp. 1415, Jan. 1999.

[4] E. Zimmermann, W. Rave, and G. Fettweis, "On the complexity of sphere decoding," in *Proc. International Symp. on Wireless Pers. Multimedia Commun.*, Abano Terme, Italy, Sept. 2004.

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

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

[7] J. Jalden and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. Signal Processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

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

[9] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Select. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.

[10] M. Siti and M. Fitz, "A novel soft-output layered orthogonal lattice detector for multiple antenna communications," in *Proc. IEEE ICC*, vol. 4, pp. 1686–1691, June 2006.

[11] T. Cui and C. Tellambura, "Joint data detection and channel estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 54, pp. 670–679, Apr. 2006.

[12] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2389–2402, Oct. 2003.

[13] C. P. Schnorr and M. Euchner, "Lattice basis reduction: improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, 1994, pp. 181–191.

[14] D. Wubben, R. Bohnke, J. Rinas, V. Kuhn, and K. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *Electron. Lett.*, vol. 37, no. 22, pp. 1348–1350, Oct. 2001.

[15] J. Anderson and S. Mohan, "Sequential coding algorithms: a survey and cost analysis," *IEEE Trans. Commun.*, vol. 32, pp. 169–176, Feb. 1984.

[16] Z. Guo and P. Nilsson, "A VLSI architecture of the Schnorr-Euchner decoder for MIMO systems," in *Proc. IEEE CAS Symp. Emerging Technologies*, vol. 3, pp. 65–68, June 2004.

[17] K. Wong, C. Tsui, R.-K. Cheng, and W. Mow, "A VLSI architecture of a K-best lattice decoding algorithm for MIMO channels," in *Proc. IEEE ISCAS02*, vol. 3, pp. 273–276, 2002.

[18] A. Burg, M. Wenk, M. Zellweger, M. Wegmueller, N. Felber, and W. Fichtner, "VLSI implementation of the sphere decoding algorithm," in *Proc. 30th European Solid-State Circuits Conference 2004*, pp. 303–306, Sept. 2004.

[19] L. Qingwei and W. Zhongfeng, "Improved K-best sphere decoding algorithms for MIMO systems," in *Proc. IEEE ISCAS*, pp. 4, May 2006.

[20] D. Waters and J. Barry, "The Chase family of detection algorithms for multiple-input multiple-output channels," in *Proc. IEEE GLOBECOM*, vol. 4, pp. 2635–2639, Nov. 2004.

[21] T. Cui and C. Tellambura, "Generalized feedback detection for MIMO systems," in *Proc. IEEE GLOBECOM*, vol. 5, pp. 5, Dec. 2005.

[22] G. Rekaya and J. Belfiore, "On the complexity of ML lattice decoders for decoding linear full-rate space-time codes," in *Proc. 2003 IEEE International Symposium on Information Theory*, pp. 206–206, July 2003.

[23] W. Zhao and G. Giannakis, "Sphere decoding algorithms with improved radius search," *IEEE Trans. Commun.*, vol. 53, pp. 1104–1109, July 2005.

[24] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2201–2214, Aug. 2002.