Progressive Hypercube Decoding

Francisco A. Monteiro[1,2], Ian J. Wassell[2]

[1]*Department of Engineering, University of Cambridge*

[2]*Digital Technology Group, The Computer Laboratory, University of Cambridge*
15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom

{fatbnm2,ijw24}@cam.ac.uk

***Abstract*— This paper presents a detection technique for multiple-input multiple-output (MIMO) spatial multiplexing systems that is based in a quantized received lattice. This permits the search to be focused into successive ordered subspaces. The proposal greatly reduces the number Euclidian distances need to be calculated replacing it by memory usage for back tracing candidate vectors. Thus, this tool enables further pruning the search tree when using a sphere decoder. The receiver starts by quantizing both the received vector and the lattice points and then defines a neighbourhood around the cell containing the received vector. If needed, the neighbourhood is extended to a larger Manhattan distance. Conversely, dense clusters of lattice points are resolved by increasing the quantization bits per dimension. The paper describes the probability density functions (pdf) of the lattice components and presents numerical results for the pdf of the number of lattice points that are candidates to be evaluated in a subsequent block based on Euclidean distances. All the results are for the frequency flat fast fading channel.**

I. INTRODUCTION

In many wireless propagation scenarios there is a sufficiently rich scattering environment giving raise to multipath. This phenomenon can be exploited in order to increase the diversity of the link and consequently reduce the symbol error rate (SER), or, on the other hand, increase the overall data rate. The later type of MIMO systems [1], perform spatial multiplexing by transmitting several data streams in parallel and therefore increasing the spectral efficiency of the physical layer. However, this is achieved at the expense of an increasing complexity (in terms of the number of operations) that grows exponentially with the number of transmission antennas. The impact of the number of receive antennas is much less dramatic, only causing a linear increase of the memory and number of operations at the receiver. Some linear and non-linear receivers are well known today [1], [2], [3]. The simplest ones, i.e., the linear ones, use the zero-forcing (ZF) or the minimum mean square error (MMSE) criteria. The most important examples of non-linear receivers are the vertical Bell Laboratories layered space-time receiver (BLAST), also called the ordered successive interference cancellation (OSIC) receiver [2], [3], and the receivers based in lattice reduction [3], [4], [5]. The OSIC receiver is sub-optimal as it is not able to extract the full diversity of the system configuration, i.e., the curves representing the number of errors against the signal to noise ratio (SNR) are less steep than the ones of the ML receiver. The lattice reduction aided receivers are able to extract from the system the same diversity as the ML detector [5] in spite of introducing some power penalty (i.e., the SER curves are parallel to the ML curves, though typically degraded by 3 to 5 dB).

Sphere decoding (SD) [6], [7] (originally devised for lattice decoding [8]) attains the performance of ML with a much smaller number of operations [2], [3], and has inspired several derivations, e.g., [9], [10], [11]. For N-dimensional hyperspheres, the number of operations is $O(N^2)$ to $O(N^3)$ in the high SNR regime; at low SNR it still requires a exponential number of operations when N grows linearly [9] (details in [2]). The idea behind sphere decoding is that only a reduced number of signals which yield a received point inside a hypersphere centred in the received signal will be considered instead of performing an exhaustive search over all possible combinations of transmitted of symbols. The complexity (and consequently the speed) of the SD algorithm is very dependent on the radius of the considered sphere (which value is selected heuristically [2], [3], and gives raise to some problems regarding the ordering of the searched points [10] (and references therein).

Many different metrics to quantify the detection complexity can be found in the literature; for example, complexity can be measured by the number of flops required [7], by estimating the average number of expanded nodes [10], by comparing the computation time of the algorithms in a computer [9] or by counting the total number of multiplications required by the detection algorithm [11].

Some proposals have been made in order implement multiplication free receivers either by using approximated expressions for the squared Euclidean distances SED [12] or by means of look-up tables [13], [14], which are only possible use to use after quantizing the problem. In this paper we propose a technique that takes further advantage of the quantized space by performing SD over the quantized lattice points, which is a much smaller number than the original number of points in the lattice noticing that clusters of points collapse into the same *cell* (or *hypercube*).

II. LINEAR, CANCELLATION AND EXACT DETECTORS

A MIMO system with flat Raleigh fading in each one of its sub-channels has the following complex baseband model

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

where $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_T}]^T$, that is, each component is transmitted from each one of the N_T transmit antennas, $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_R}]^T$, each component corresponding to the signal in each one of the N_R receiver antennas and the noise vector $\mathbf{n}_g = \left[\mathbf{n}_{g,1}, \mathbf{n}_{g,2}, \ldots, \mathbf{n}_{g,N_R}\right]^T$ is composed of complex Gaussian circularly symmetric random variables with zero mean and variance $\sigma_{n_g}^2$ (i.e., $\sigma_{n_g}^2 / 2$ in both real and imaginary parts). The channel matrix $\mathbf{H}$ has dimensions $N_R \times N_T$ system is all its $h_{i,j}$ elements i.i.d complex circularly symmetric Gaussian random variables with zero mean and unit variance (i.e., with variance 0.5 in both the real and imaginary components), so that no gain is introduced. The components in $\mathbf{x}$ are taken from a set $\Xi_{\mathbf{C}}$ of symbols belonging to an M-ary complex constellation with real and imaginary parts taken from the real set $\Xi_{\mathbf{R}}$. The symbols are taken from a given constellation $\Xi_{\mathbf{C}}$ having an average signal power σ_x^2. Thus, for a given SNR, the power $\sigma_{n_g}^2$ is determined from

$$SNR = \frac{\underset{\mathbf{H},\mathbf{x}}{\mathrm{E}}\left\{\|\mathbf{Hx}\|^2\right\}}{\underset{\mathbf{n}}{\mathrm{E}}\left\{\|\mathbf{n}\|^2\right\}} = \frac{N_T N_R \sigma_x^2}{N_R \sigma_{n_g}^2} = \frac{N_T \sigma_x^2}{\sigma_{n_g}^2}. \qquad (2)$$

This complex valued model can be converted into one with the double number of dimensions where all the elements are real numbers by stacking the real and imaginary parts of $\mathbf{x}$ and $\mathbf{y}$ and thus constructing an equivalent real valued description

$$\begin{bmatrix} \Re\mathbf{y} \\ \Im\mathbf{y} \end{bmatrix} = \begin{bmatrix} \Re\mathbf{H} & -\Im\mathbf{H} \\ \Im\mathbf{H} & \Re\mathbf{H} \end{bmatrix} \begin{bmatrix} \Re\mathbf{x} \\ \Im\mathbf{x} \end{bmatrix} + \begin{bmatrix} \Re\mathbf{n} \\ \Im\mathbf{n} \end{bmatrix} \quad \text{or} \quad \mathbf{y}_r = \mathbf{H}_r \mathbf{x}_r + \mathbf{n}_r \qquad (3)$$

where $\Re$ and $\Im$ denote the real and imaginary components respectively. Thus $\mathbf{x}_r = \left[\mathbf{x}_{r,1}, \mathbf{x}_{r,2}, \ldots, \mathbf{x}_{r,2N_T}\right]^T$, $\mathbf{y}_r = \left[\mathbf{y}_{r,1}, \mathbf{y}_{r,2}, \ldots, \mathbf{y}_{r,2N_R}\right]^T$, and $\mathbf{n}_r = \left[\mathbf{n}_{r,1}, \mathbf{n}_{r,2}, \ldots, \mathbf{n}_{r,2N_R}\right]^T$. Also, $\mathbf{H}_r$ has double dimensions with respect to $\mathbf{H}$. The problem of detecting the transmitted symbols is optimally solved by maximum likelihood detection based on the (squared) ℓ_2 norm, i.e., the SED. The problem is

$$\begin{aligned} \hat{\mathbf{x}}_{r(ML)} &= \min_{\mathbf{x}_r \in \Xi_{\mathbf{R}}^{2N_T}} \left\{\|\mathbf{y}_r - \mathbf{H}_r \mathbf{x}_r\|^2\right\} \\ &= \min_{\mathbf{x}_r \in \Xi_{\mathbf{R}}^{2N_T}} \left\{(\mathbf{y}_r - \mathbf{H}_r \mathbf{x}_r)^T (\mathbf{y}_r - \mathbf{H}_r \mathbf{x}_r)\right\} \end{aligned} \qquad (4)$$

This implies the measurement and comparison of M^{N_T} squared Euclidian distances per component of the $\mathbf{R}^{2N_R}$ space or per component of the complex space $\mathbf{C}^{N_R}$ in (1). Hence, the search increases exponentially with the number of transmission antennas for a given modulation.

The minimization in (4) is equivalent to

$$\hat{\mathbf{x}}_{r(ML)} = \min_{\mathbf{x}_r \in \Xi_{\mathbf{R}}^{2N_T}} \left\{(\mathbf{x}_{rZF} - \mathbf{x}_r)^T \mathbf{H}_r^T \mathbf{H}_r (\mathbf{x}_{rZF} - \mathbf{x}_r)\right\} \qquad (5)$$

where $\mathbf{x}_{\mathrm{ZF}}$ is the zero-forcing (ZF) solution that is obtained by the simplest linear receiver, i.e., the one that simply inverts the channel matrix. In the general case, as the channel matrix $\mathbf{H}$ (nor $\mathbf{H}_r$) is not necessarily square (corresponding to $N_T = N_R$), then the Moore-Penrose pseudo-inverse of $\mathbf{H}$ is used, which is given by $\mathbf{H}^+ = \left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H$ where $(\cdot)^H$ denotes the Hermitian transposition. In this case the "filtering" matrix is $\mathbf{W}_{ZF} = \mathbf{H}^+$ and therefore $\mathbf{x}_{\mathrm{ZF}} = \mathrm{Slice}\left[\mathbf{H}^+ \mathbf{y}\right]$, where the "Slice" operation is used to detect the symbols in a constellation (the same applies to $\mathbf{x}_{r\mathrm{ZF}}$). This procedure induces noise enhancement in the constellation space where the decisions of constellation symbols are to be made. The MMSE receiver takes the noise into account, resulting in the following "filtering" matrix $\mathbf{W}_{MMSE} = \left(\mathbf{H}^H \mathbf{H} + \frac{1}{SNR} \mathbf{I}_{n_T}\right)^{-1} \mathbf{H}^H$.

The OSIC-ZF receiver uses $\mathbf{W}_{ZF}$ in the initial iteration. The component of y_r with the smallest noise amplification by the corresponding row of $\mathbf{W}_{\mathrm{ZF}}$ is selected and detected. The next step is to remodulate that symbol and subtract its effect from the original received $\mathbf{y}_r$. This procedure is repeated for the new vector, originating the detection of a second component of $\mathbf{y}$, and is repeated until all components have been detected. The OSIC-MMSE receiver operates similarly by applying $\mathbf{W}_{MMSE}$ instead of $\mathbf{W}_{ZF}$. The complexity of OSIC is only polynomial with N_T [10] but the fact that an erroneous decision in one component cascades the error throughout the components to be subsequently detected explains why OSIC does not entirely exploit the diversity available in the system.

The SD is an exact algorithm that makes use of (5) and bounds the maximum distance (radius) for the lattice points that are to be searched around $\mathbf{y}_{\mathrm{r}}$ by imposing

$$(\mathbf{x}_{r(ZF)} - \mathbf{x}_r)^T \mathbf{H}_r^T \mathbf{H}_r (\mathbf{x}_{r(ZF)} - \mathbf{x}_r) < C. \qquad (6)$$

By means of a factorization that can decompose $\mathbf{H}^T\mathbf{H}$ into a unitary matrix and an upper triangular matrix $\mathbf{R}$ (such as the Cholesky -when $\mathbf{H}_r$ is Hermitean- or the QR factorization), the inequality (6) can be re-written as

$$\sum_{i=1}^{N_T} \left(\sum_{j=1}^{N_T} \mathbf{R}_{i,j} \left(x_{r(ZF),j} - x_{r,j}\right) \right)^2 \leq C, \qquad (7)$$

noticing that the unitary matrix will not affect the norm.

This allows us to compute (4) in a progressive manner, including the components of $\mathbf{x}$ one at the time. Thus, once the inequality does not hold for a certain symbol x_i, there is no need to try the inequality including higher dimensions in the sum (i.e., other components of $\mathbf{x}$). Whenever the last dimension is reached one has a combination forming $\mathbf{x}$ that obeys the inequality and the next step is to reduce the radius C for a subsequent search.

III. Uniform Multidimensional Cells

The detection technique presented in this paper operates in a space where both the lattice points and the received signal are uniformly quantized component-wise, giving rise to a slicing of the received space into uniform hypercubes (or cells). The clipping value is selected from the maximum component taken over all lattice points, i.e., $y_{i,sat} = \max\left\{y_{r,i}^{(l)}\right\}$. Both the received signal and the lattice itself are bounded to

$[-y_{i,sat}, +y_{i,sat}]$ in the i^{th} real dimension, corresponding to the clipping imposed by $Q(\cdot)$, This creates a closed space that is an hypercube with edges of size $2y_{sat}$. By doing this no lattice point will suffer any quantization error due to saturation, however, if $\mathbf{y}_r$ lies outside that space, it will be pulled (projected) into the closet face of the hypercube. It can be shown that this does not introduce any detection error [14].

According to (3), each component in $\mathbf{y}_r$ is of the form

$$\begin{cases} y_{r,i} = \Re h_{i,1}\cdot\Re x_1 + \cdots + \Re h_{i,N_T}\cdot\Re x_{N_T} \\ \quad -\Im h_{i,1}\cdot\Im x_1 - \cdots - \Im h_{i,N_T}\cdot\Im x_{N_T} + \Re n_i, \quad i\le N_R \\ y_{r,i} = \Im h_{i-N_R+1,1}\cdot\Re x_1 + \cdots + \Im h_{i-N_R+1,N_T}\cdot\Re x_{N_T} \\ \quad +\Re h_{i-N_R+1,1}\cdot\Im x_1 + \cdots + \Re h_{i-N_R+1,N_T}\cdot\Im x_{N_T} + \Im n_{i-N_R+1}, \\ \qquad\qquad N_R < i \le 2N_R \end{cases} \tag{8}$$

In the particular case of QPSK, both $\Re x_i$ and $\Im x_i$ are taken from $\Xi_{\mathbf{R}}^{2N_T}=\{-1,+1\}$ and therefore the probability for the symbols components is $P(\Re x_i = -1) = P(\Re x_i = +1) = 0.5$ and $P(\Im x_i = -1) = P(\Im x_i = +1) = 0.5$. Denoting a Gaussian distribution with mean μ and variance σ^2 by $N(\mu , \sigma^2)$, then as defined in Section II, $\Re h_{i,j}\sim N(0 , 0.5)$ and $\Im h_{i,j}\sim N(0 , 0.5)$. Therefore, regardless the signal polarity affecting the Gaussian variables in (8), all the terms have a pdf $N(0 , 0.5)$ and each component of $\mathbf{y}_r$ has a pdf given by the convolution $N_1(0,0.5)\otimes N_2(0,0.5)\cdots\otimes N_{2N_T}(0,0.5)\otimes\ N_{2N_T}(0,\sigma_{n_g}^2/2)$ (the last term coming from the noise), and so $y_{r,i}\sim N(0,2N_T/2+\sigma_{n_g}^2/2)=N(0,N_T+\sigma_{n_g}^2/2)$.

In brief, the received lattice components corrupted by noise have an overall Gaussian distribution and lattice points are distributed in the receive space according to a $2N_r$-dimensional Gaussian pdf with i.i.d. components. Hence, a uniform quantization yields a sub-optimal quantization of the lattice. A comparison between uniform and optimal non-uniform quantization for Gaussian processes can be found in [15-Sec.5.3] in terms of the signal to quantization noise rations as a function of the number of quantization bits b. The gain is of the order of 2 dB for b=6 but, for the cases that we will be of interest for our proposal, say b<5, the gain is <1 dB and is negligible for b<3.

Initially, the imaginary components of $\mathbf{y}$ are stacked under the real ones, generating $\mathbf{y}_r$, as in (3). Then the received $2N_R$-dimensional space is quantized and all the possible points on the lattice are mapped into the new quantized space. Notice that these correspondences are needed to maintain in memory as they will be used to back trace the surviving candidates for the Euclidean distance test that follows. The possible points in the lattice are denoted as $\mathbf{y}_r^{(l)}=\mathbf{H}_r\mathbf{x}_r^{(l)}$ l=1, 2,…, M^{N_T} and the quantization function by $Q(\cdot)$. The resulting quantized received vector is $\tilde{\mathbf{y}}_r$ = $\left[\tilde{y}_{r,1},\tilde{y}_{r,2},\cdots,\tilde{y}_{r,2N_T}\right]^T$ = $Q\left(\left[y_{r,1},y_{r,2},\cdots,y_{r,2N_T}\right]^T\right)$. Each one of these components $\tilde{y}_{r,i}\in\{c_1,c_2,c_3,\cdots,c_L,\}$, i.e., the $L=2^b$ possible quantization levels (described by b bits) with a uniform step size $q=c_m-c_{m+1}$, for $m\in\{1,2,\cdots,L\}$. We define the quantization step as q=2 which defines the quantization levels as $\{c_1,c_2,c_3,\cdots,c_L,\}=\{-(L-1),\cdots,-3,-1,+1,+3,\cdots,(L-1)\}$. The *total quantization noise* in $\tilde{\mathbf{y}}_r$ was quantified in [14]).

It should be noticed that the distance considered so far is the squared *Euclidean distance*, i.e., the ℓ_2 norm. One of the central ideas behind the proposal in this paper is that the neighbourhood surrounding the cell containing the received vector will be defined by a ℓ_1 norm instead, also known as the squared *Manhattam distance* (SMD).

Defining the component $\Delta_i^{(l)}=\left(\tilde{y}_{r,i}-\tilde{y}_{r,i}^{(l)}\right)$ for a particular point $\tilde{\mathbf{y}}_r^{(l)}$ of the quantized lattice, then the SED from the received quantized point to each one of the quantized lattice points are

$$\left(d_{\ell_2}^{(l)}(\tilde{\mathbf{y}}_r,\tilde{\mathbf{y}}_r^{(l)})\right)^2=\sum_{i=1}^{N_R}\left(\tilde{y}_{r,i}-\tilde{y}_{r,i}^{(l)}\right)^2=\sum_{i=1}^{N_R}\left(\Delta_i^{(l)}\right)^2, \tag{9}$$

with l=1, 2, ···, M^{N_T}, whilst the corresponding SMD is

$$\left(d_{\ell_1}^{(l)}(\tilde{\mathbf{y}}_r,\tilde{\mathbf{y}}_r^{(l)})\right)^2=\sum_{i=1}^{N_R}\left|\Delta_i^{(l)}\right|,\quad l=1,2,\cdots,M^{N_T}. \tag{10}$$

IV. Progressive Neighbourhoods

The quantized received vector $\tilde{\mathbf{y}}_r$ lies inside a hypercube with sides having size q. Inside that same cell one can find any number of quantized points of the lattice, depending on the channel realization (Fig. 1). The minimum number of bits per dimension that originates a one to one correspondence of the lattice points with the hypercube was investigated in [13], [14] for different MIMO systems. The main objective for the proposed hypercube decoding is *not* to have a one to one mapping into the quantized lattice but rather to capture several lattice points inside distinct hypercubes so that some lattice points will exist inside the same hypercube where $\tilde{\mathbf{y}}_r$ lies. If no point is found inside that hypercube then it is necessary to search in the cells in contact with it. Therefore it is important to know how to generate that neighbourhood and to be able to count the number of cells in it. To do this, one should take into account the fact that closest point to $\mathbf{y}_r$ is not necessarily inside one of the hypercubes that are closer to $\tilde{\mathbf{y}}_r$, as is exemplified in Fig. 2. It is possible to see there a situation where the closest lattice point to the $\mathbf{y}_r$ is inside cell 4. However, after quantization, the selected (closest) point would be the one lying in the cell associated with $\mathbf{y}_r^{(2)}$. This would happen for any point in the shaded region (i.e., the region where the points are more distant to $\mathbf{y}_r$ than to $\mathbf{y}_r^{(4)}$ but are quantized in the cell associated with $\mathbf{y}_r^{(2)}$, because it would be closer to $\tilde{\mathbf{y}}_r$). Because of the existence of cases such as this in Fig. 2, the neighbourhood should include all the cells in contact with the central hypercube. That neighbourhood is defined by the set of cells $N_c^{(n)}(\tilde{\mathbf{y}}_r) = \{\tilde{\mathbf{y}}_r + \mathbf{\Delta}\}$ for all possible $\mathbf{\Delta} = [\Delta_1, \Delta_2, \ldots, \Delta_{2N_R}]$ with the Δ_i components taken from $\{-n, \ldots, -1, 0, +1, \ldots, +n\}$, and n is natural number. Note that $N_c^{(n)}(\tilde{\mathbf{y}}_r)$ obeys $d_{\ell_2}^2(\tilde{\mathbf{y}}_r + \mathbf{\Delta}, \tilde{\mathbf{y}}_r) \le C$, with $C = \sum_{i=1}^{2N_R} (n \cdot q)^2$.

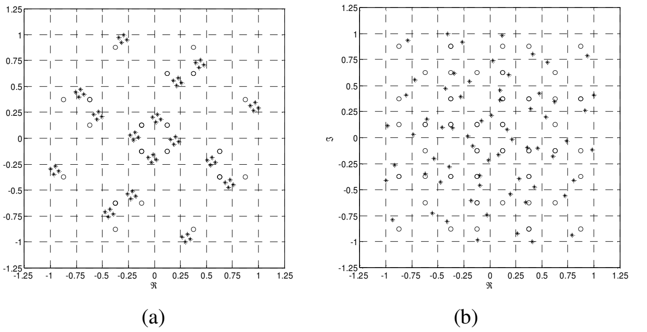

Fig 1. Quantized lattice with L=8 (white circles) of the original complex lattice (black dots) in the first receive antenna of a 2 ×2 system with QPSK symbols ($M^{N_T} = 16$) for two different channel realizations.

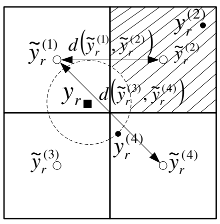

Fig. 2. Region associated to a particular 2D case where deciding according to the distance among cells leads to a wrong decision as the point in cell 2 is perceived to be closer to the point in cell 1 instead of the correct one in cell 4.

Notice that the cells in the neighbourhood $N_c^{(n)}(\tilde{\mathbf{y}}_r)$ can even be easily ordered by construction, running $\mathbf{\Delta}$ from $[-n, -n, \ldots, -n]_{(1\times 2N_R)}$ to $[+n, +n, \ldots, +n]_{(1\times 2N_R)}$.

The number of neighbouring cells for a given n is $(2n+1)^{2N_R} - 1$ (excluding the central cell itself). By definition the central cell is its own zero-order neighbourhood $N_c^{(0)}(\tilde{\mathbf{y}}_r)$. For the first-order neighbourhood, $N_c^{(1)}(\tilde{\mathbf{y}}_r)$, and for the simple case of 2 dimensions ($2N_R = 2$) one can see that the components Δ_i are drawn from $\{-1, 0, +1, \}$ and therefore the number of combinations for $\mathbf{\Delta}$ gives $|N_c^{(1)}(\tilde{\mathbf{y}}_r)| = 3^2 - 1 = 8$ neighbours; in 3 dimensions the number grows to $|N_c^{(1)}(\tilde{\mathbf{y}}_r)| = 3^3 - 1 = 26$ neighbours. Both cases are depicted in Fig. 3. Despite being simple to visualize, neither of these cases correspond to a possible MIMO system (the first corresponds to a single receive antenna and the number of dimensions, $2N_R$, must be even). For a 2 ×2 system $|N_c^{(1)}(\tilde{\mathbf{y}}_r)| = 3^4 - 1 = 80$ and for a 4 ×4 system that number grows to 6560 cells.

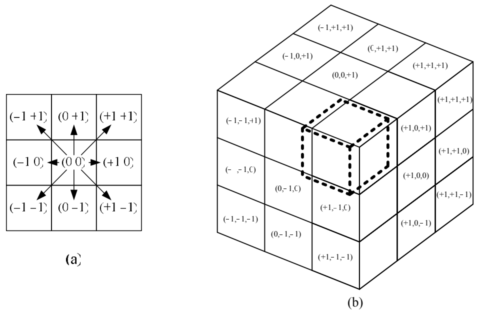

Fig. 3. First-order neighbours of a cell according to the Euclidian distance. (a) 8 neighbours in 2 dimensions. (b) 26 neighbours in 3 dimensions. Some displacement vectors $\mathbf{\Delta}$ are indicated.

V. THE ALGORITHM

Let us consider the constrained detection for a defined n

$$\hat{\mathbf{x}}_{r(ML)} = \min_{\tilde{\mathbf{y}}_r^{(l)} \in N_c^{(n)}(\tilde{\mathbf{y}}_r)} \left\{ \left\| \tilde{\mathbf{y}}_r - \tilde{\mathbf{y}}_r^{(l)} \right\|^2 \right\}. \tag{11}$$

The selection of the points to be included in $N_c^{(n)}(\tilde{\mathbf{y}}_r)$ would require to check all the M^{N_T} lattice points, involving the same number of calculated distances in ML decoding. The benefit of dealing with an integer problem is that one can perform the neighbourhood test in an easier way. As the number of neighbours increases exponentially with the number of dimension it is desirable to select the n^{th}-order neighbourhood with respect to the ℓ_1 norm instead of ℓ_2 and thus eliminate all the multiplications. This simplification does not lead to any errors as at this stage the receiver only selects a (reduced) number of candidate vectors for subsequent detection via minimum SED (which will always include the best vector). Fig. 4 depicts the difference between the first-order neighbourhoods defined by norms ℓ_2 and ℓ_1 (in 2D).

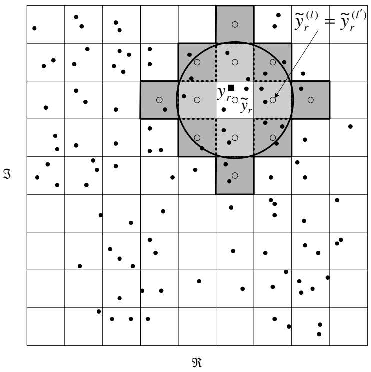

Fig. 4. First-order (light shading) and second-order (darker shading) neighbourhoods according to the Manhattan distance in a bidimensional lattice (i.e., N_R=1). The circle indicates the first-order ℓ_2 neighbourhood.

By evaluating if $d_{\ell_1}(\tilde{\mathbf{y}}_r - \tilde{\mathbf{y}}_r^{(l)}) \le n$, for l=1, 2, ···, M^{N_T}, it is possible to determine which lattice points belong to the respective n^{th}-order neighbourhood $N_c^{(n)}(\tilde{\mathbf{y}}_r)$.

The detection starts by considering a pre-defined neighbourhood. If no lattice points were found, n increases until at least one point is found in the vicinity of $\tilde{\mathbf{y}}_r$. Conversely, if a dense cluster of lattice points is found at any stage, i.e, the number of points with $d_{\ell_1}\left(\tilde{\mathbf{y}}_r - \tilde{\mathbf{y}}_r^{(l)}\right) = 0$ is higher than a certain threshold, then one increases the grid resolution by 1 bit.

VI. Results and Conclusions

Considering the analysis in Section III, the pdf for the non-quantized lattice is $\left(y_{r,i} - y_{r,i}^{(l)}\right) \sim 2N(0, N_T + \sigma_{n_g}^2/2) = N(0, 2N_T + \sigma_{n_g}^2)$. For the terms in sum of the SMD one has $\left|\left(y_{r,i} - y_{r,i}^{(l)}\right)\right| \sim 2N(0, 2N_T + \sigma_{n_g}^2)u(x) \equiv g(x)$ (see Fig. 5), where $u(x)$ is the Heaviside step function (1 for x>0 and 0 for x<0). Finally, given the sum in (10), the pdf for the SMD will be the convolution $g_1(x) \otimes g_2(x) \cdots \otimes g_{2N_R}(x)$, which is difficult to obtain analytically. Fig. 5 also shows the final result for this convolution and therefore for the distribution of the SMD. Note that for the equivalent pdf for the SED one has a chi-squared-like distribution with $2N_R$ degrees of freedom (constructed from Gaussians with non unitary variance).

Fig. 6 depicts the pdf for the number of lattice points inside two different neighbourhoods (defined by normalized Manhattan distances in the quantized lattice). In both cases the average is much smaller than the original 256 points (and it can be further reduced). Fig. 7 shows that the performance attained with the specifications of Fig.6(b) is within 3 dB of the ML curve.

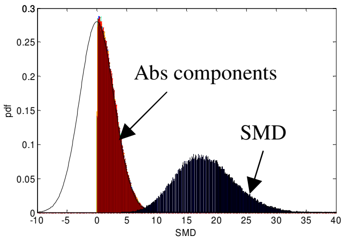

Fig. 5. Pdf for the abs components (ligher) of the SMD given by (10) and pdf of the final SMD (darker) obtained from the $2N_R$ convolutions of the pdf's of the abs components (for N_R=4 and SNR=15dB)

References

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar and H. Bölcskei, "An overview of MIMO communications - a key to gigabit wireless", *Proc. of the IEEE*, vol. 92, No. 2, pp. 198–218, Feb. 2004.

[2] T. Kailath, H. Vikalo, B. Hassibi, H. Bölcskei (Editor), D. Gesbert (Editor), C. B. Papadias (Editor), A.-J. van der Veen (Editor), "MIMO receive algorithms", in *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, Chapter 15, Cambridge University Press, 2006.

[3] C. Windpassinger, "Detection and precoding for multiple input multiple output channels", Ph.D. thesis, University of Erlangen-Nurenberg, Nuremberg, Germany, 2004.

[4] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems, in *Proc. of. IEEE GLOBECOM'02 – 2002 IEEE Global Telecom. Conf.*, Vol. 1, Taipei, Taiwan, pp. 424–428, Nov. 2002.

[5] M. Taherzadeh, A. Mobasher and A. K. Khandani, "LLL Lattice-Basis Reduction Achieves Maximum Diversity in MIMO Systems", in *Proc. of ISIT'05 – IEEE Int. Symp. on Information Theory*, Adelaide, Australia, Sept, pp. 1300–1304, 2005.

[6] E. Agrell, T. Eriksoon. A. Vardy and K. Zeger. "Closest point search in lattices", *IEEE Trans. on Information Theory*, Vol. 48, pp. 2201–2214, Aug. 2002.

[7] Ami Wiesel, Xavier Mestre, Alba Pagés, Javier R. Fonollosa, "Efficient implementation of sphere decoding" in *Proc of SPAWC'03 – 4th IEEE Workshop on Signal Processing Advances in Wireless Communications*, Rome, Italy, June 2003, pp. 18–21.

[8] E. Viterbo and J. Boutros, "A Universal Lattice Code Decoder for Fading Channels", *IEEE Transactions on Information Theory*, Vol. 45, No. 5, pp. 1638–1642, July 1999.

[9] M. Alghoniemy and A. H. Tewfik, "Cube decoding", *in Proc. of ICC'05 - The IEEE Int. Conf. on Communications*, Vol. 3, Seoul, Korea, May 2005, pp. 1901–1905.

[10] K. Su and I. Wassell, "A new ordering for efficient sphere decoding", in *Proc. of ICC'05 - The IEEE Int. Conf. on Communications*, Vol. 3, Seoul, Korea, May 2005, pp. 1906–1910.

[11] H. Sung, J. W. Kang, K. B. Lee, "A simple maximum likelihood detection for MIMO systems", *IEICE Trans. on Comm.*, Vol. E89-B, No. 8, pp. 2241–2244, August 2006.

[12] M. Rupp, G. Gritsch, H. Weinrichter, "Approximate ML detection for MIMO systems with very low complexity", in *Proc. of ICASSP' 04 – The 2004 IEEE Intern. Conf. on Acoustics, Speech, and Signal Processing*, Vol. IV, pp.809–812, Montreal, Canada, 2004.

[13] F. Monteiro, I. Wassell, "Euclidean Distances in Quantized Spaces with Pre-stored Components for MIMO Detection", in *Proc. of ECWT' 07 – 10th European Conf. on Wireless Technology – European Microwave Week*, Munich, Germany, Oct. 2007.

[14] F. Monteiro, I. Wassell, "Efficient Scalar Quantization for MIMO Spatial Multiplexing Receivers", in *Proc. of ISCTA' 07 – 9th Inter. Symp. on Communication Theory and Applications*, Ambleside, Lake District, UK, July 2007.

[15] N. Benvenuto, U. Cherubini, *Algorithms for Communications Systems and their Applications*, Willey, 2002.

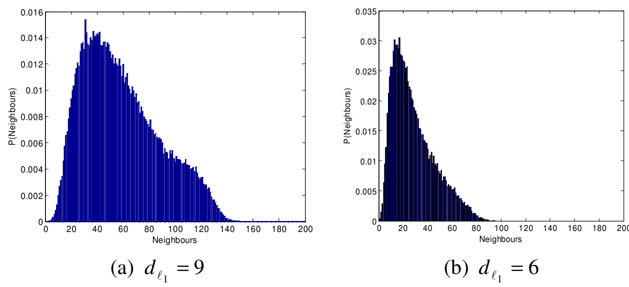

Fig. 6. Probability for the number of lattice points inside a region defined by two different normalized Manhatan distances. 2×2 system with QPSK, i.e., with $M^{N_T} = 256$; and b= 2 bits (L=4 levels) per dimension; SNR=15dB.

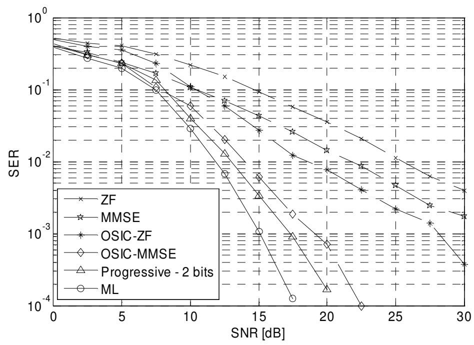

Fig. 7. SER versus average SNR for standard receivers and progressive hypercube detection with a neighbourhood defined by a normalized Manhatan distance 6 in a 4×4 system using QPSK ($M^{N_T} = 256$).