Regular Article

Quantization Error Correction Schemes for Lattice-Reduction Aided MIMO Detectors

Tien Duc Nguyen[1], Tadashi Fujino[1], Xuan Nam Tran[2]

[1]Dept. Information and Communication Engineering, University of Electro-Communications, Tokyo, Japan
[2]Dept. Communications Engineering, Le Quy Don Technical University, Hanoi, Vietnam

Correspondence: Xuan Nam Tran, namtx@mta.edu.vn

Abstract– Lattice reduction aided (LRA) linear detectors have been known to achieve near optimal performance at low complexity. However, one weakness of LRA detector is that the quantization step in LRA detector is not optimal. Based on simulation results, we show that most of detection errors in LRA linear detectors are due to quantization errors. We then propose two methods to correct the quantization errors. In the first method, sphere detectors are introduced to correct quantization errors at low additional complexity. As a second approach, we propose a list quantization scheme which can generate a list of candidate symbols from the original LRA estimated symbols. From these listed symbols, decisions are made according to the minimum Euclidean distance between the received and estimated points. It is shown by simulations that both methods provide significant BER performance improvements with only a small additional complexity.

Keywords– ZF, MMSE, lattice reduction, MIMO, spatial multiplexing, quantization error, sphere detection.

1 Introduction

Multiple-input multiple-output (MIMO) wireless systems were shown to achieve enormous channel capacity in rich-scattering wireless environments [1]. As one of the practical approaches to achieve the MIMO capacity, spatial multiplexing was demonstrated to have spectral efficiencies of 20–40 b/s/Hz in an indoor environment using the V-BLAST (Vertical-Bell Labs Layered Space-Time) detector [2]. It is known that V-BLAST is a promising detector in terms of balancing its BER performance and associated computational complexity. The BER performance of V-BLAST detectors is far from that of the optimum detector using maximum likelihood (ML) criterion.

Recently, there has been a great interest in seeking detectors with good BER performance but low computational complexity. One effective approach is to apply lattice reduction into the MIMO channel matrix to make it closer to orthogonal before using linear detectors such as zero forcing (ZF) or minimum mean square error (MMSE) to estimate transmitted symbols. This kind of detectors is often known as lattice reduction aided (LRA) linear detectors [3, 4]. It is interestingly noted from the literature that the LRA linear detectors allow to achieve the same diversity order of the ML detector but with some loss in average bit energy per noise variance (E_b/N_0).

We note that in the previous proposed LRA detectors, a lattice reduction algorithm, such as the well-known Lenstra, Lenstra and Lovász (LLL) [5], is often used to transform the original signal constellation into a lattice of parallel points. In fact, this transformation makes the optimal quantization too complex, and thus a suboptimal unconstrained element-wise quantization is used in previous works [3, 6]. The effect of the suboptimal quantization is a degradation in BER performance as mentioned above.

The objective of this paper is to find methods to save this degraded E_b/N_0 in LRA linear detectors. Using experiments by computer simulation we found that most of detection errors resulted from LRA linear detectors are due to quantization errors. Based on this observation, we propose two methods to correct the quantization errors. In the first method, sphere detectors are introduced with the aim to correct the quantization errors at low additional complexity. As a second approach, we propose a list quantization scheme which can generate a list of candidate symbols from the original LRA estimated symbols. From these listed symbols, decisions are made according to the minimum Euclidean distance between the received and estimated points. It is shown by simulation that both methods provide significant BER performance improvements at the cost of only a small additional complexity, particularly, in high E_b/N_0 region.
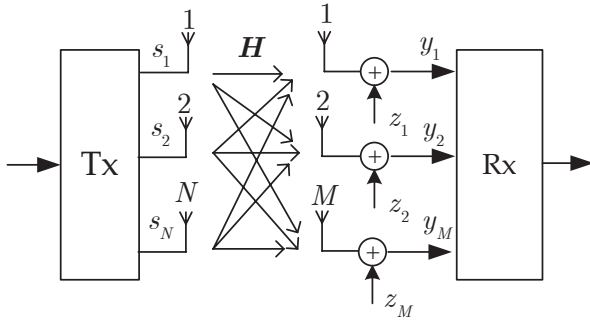
Figure 1.   System model of a MIMO-SDM system.

The remainder of the paper is organized as follows. In Section 2, we introduce the signal model and conventional linear detectors (namely ZF and MMSE) for MIMO systems. The principle of lattice reduction and its application in linear detectors are presented in Section 3. Explanation on the quantization error of the LRA-linear detectors is given in Section 4. Combined Sphere Detector (SD)-LRA detector is explained in Section 4. List quantization scheme for LRA detection is proposed in Section 5. Simulation results are shown in Section 6 and, finally, the paper is summarized and concluded in Section 7.

2 Linear MIMO Detectors

2.1 System Model

We consider a MIMO system as illustrated in Figure 1. The transmitter is equipped with N, and the receiver M antenna elements. It is also assumed that $N \leq M$. The wireless propagation environment between the transmitter and the receiver is assumed rich scattering and flat in frequency. This assumption allows to model the complex channel gains h_{mn} between the nth transmit antenna and the mth receive antenna as statistically independent Gaussian random variables with zero mean and unit variance, i.e., $h_{mn} \sim \mathcal{N}_c(0,1)$, $n = 1, 2, \ldots, N$, and $m = 1, 2, \ldots, M$. Further, the noise samples z_m at the mth receive antenna are modeled as independent complex Gaussian random variables with zero mean and variance σ_z^2, i.e., $z_m \sim \mathcal{N}_c(0, \sigma_z^2)$. The vector signal model of the system under consideration can be then expressed as

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

where $y = [y_1, y_2, \ldots, y_M]^T$ is the receive signal vector,

$$ H = \begin{bmatrix} h_{11} & h_{12} & \ldots & h_{1N} \\ h_{21} & h_{22} & \ldots & h_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1} & h_{M2} & \ldots & h_{MN} \end{bmatrix} $$
$$ = [h_1, h_2, \ldots, h_N] $$

the $M \times N$ MIMO channel matrix, $s = [s_1, s_2, \ldots, s_N]^T$ the transmit signal vector, and $z = [z_1, z_2, \ldots, z_M]^T$ the noise vector. For simplicity, we make the further assumptions about the transmit power constraint, noise

and signal correlation are as follows: $E\{ss^H\} = \zeta_s^2 I_N$, and $E\{sz^H\} = 0$, where 0 denotes a zero matrix of appropriate size, respectively. In this paper, $E\{\cdot\}$ represents the expectation and $(\cdot)^H$ the complex transpose operation.

2.2 Linear Detectors

In the linear detectors, the receive signal y is multiplied with a weight matrix W to get the estimates of the transmitted symbols

$$ \tilde{s} = W^H y. \tag{2} $$

The estimated vector $\tilde{s}$ is then decided to the nearest points in the signal constellation.

Based on the ZF and MMSE methods, the weight matrices of the corresponding detectors are given by [7]

$$ W_{ZF} = H^\dagger = \left(H^H H\right)^{-1} H^H, \tag{3} $$
$$ W_{MMSE} = (H^H H + \rho I_N)^{-1} H^H, \tag{4} $$

where $H^\dagger$ denotes the pseudo-inverse of H, and $\rho = \sigma_z^2 / \zeta_s^2$ represents the inverse signal to noise ratio (SNR) at the input of the receiver.

Using (2) and (3), the estimate of the transmitted vector s using the ZF detector is given by

$$ \tilde{s}_{ZF} = W_{ZF} y = H^\dagger y = s + H^\dagger z. \tag{5} $$

Each element of $\tilde{s}_{ZF}$ is then mapped to a signal constellation point to make the final decision on the detected symbols.

For the MMSE detector, we can use the ZF equivalent representation proposed by Hassibi [8] by defining the following extended matrices:

$$ \bar{H} = \begin{bmatrix} H \\ \sqrt{\rho} I_N \end{bmatrix} \quad \text{and} \quad \bar{y} = \begin{bmatrix} y \\ 0 \end{bmatrix}. \tag{6} $$

Similar to the ZF detector, the estimate of the transmitted vector s using the MMSE detector can be then given by

$$ \tilde{s}_{MMSE} = W_{MMSE} y = \bar{H}^\dagger \bar{y}. \tag{7} $$

Due to this similarity in the expression between the MMSE and ZF detectors, analysis for the ZF detector will also be applied for the MMSE detector in the following parts.

In the next section, we will briefly introduce the application of lattice reduction to the linear detectors in order to improve the BER performance.

3 Lattice Reduction and LRA Linear Detectors

3.1 Principle of Lattice Reduction

The basic concept of lattice reduction [5] is to consider the noiseless receive signal vector in (1),

$$ \mathcal{L} = Hs = s_1 h_1 + s_2 h_2 + \ldots + s_N h_N \in \mathbb{C}^M, $$
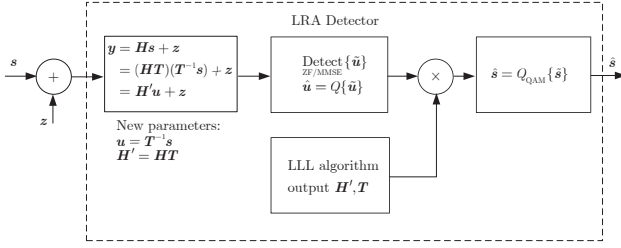
Figure 2.   LRA detectors block diagram.

as a (complex) lattice spanned by the generator (channel) matrix $\boldsymbol{H}$. The set of column vectors $\{\boldsymbol{h}_1, \boldsymbol{h}_2, ..., \boldsymbol{h}_N\}$ is called a basis of lattice $\mathcal{L}$. A lattice may have many different bases.

The aim of lattice basis reduction, or often simply referred to as lattice reduction (LR), is to transform a given basis $\boldsymbol{H}$ into a new basis $\boldsymbol{H}'$ whose columns are close to mutually orthogonal. The detection will be then carried out based on $\boldsymbol{H}'$, followed by a quantization operation. The most popular algorithm for LR is the LLL algorithm [5]. In fact, the whole LR algorithm can be described by a linear transform $\boldsymbol{H}' = \boldsymbol{HT}$, where $\boldsymbol{H}'$ is the LLL-reduced matrix. The transform matrix $\boldsymbol{T}$ is an integer unimodular matrix with determinant $\det\{\boldsymbol{T}\} = \pm 1$, which is resulted from the LLL algorithm. Detailed steps of the algorithm are presented in [7]. Further information can be found in [4–6, 9].

3.2 LRA Detectors

The motivation of applying LR to linear detectors is to improve the decision region [3]. Using the transform matrix $\boldsymbol{T}$ obtained from a LR algorithm as described in Sect. 3.1, the system model (1) can be rewritten as [7]

$$\boldsymbol{y} = \boldsymbol{Hs} + \boldsymbol{z} = (\boldsymbol{HT})(\boldsymbol{T}^{-1}\boldsymbol{s}) + \boldsymbol{z} \equiv \boldsymbol{H}'\boldsymbol{u} + \boldsymbol{z}, \quad (8)$$

where $\boldsymbol{H}' \triangleq \boldsymbol{HT}$ and $\boldsymbol{u} \triangleq \boldsymbol{T}^{-1}\boldsymbol{s}$ are the new complex channel matrix and transmitted signal vector, respectively. Based on this new model, LRA-ZF or LRA-MMSE detector estimates $\boldsymbol{u}$ and then uses the transform matrix $\boldsymbol{T}$ to convert $\boldsymbol{u}$ to $\tilde{\boldsymbol{s}}$. Since columns of $\boldsymbol{H}'$ are closely orthogonal, the linear detection based on $\boldsymbol{H}'$ provides better BER performance [4]. A block diagram of LRA linear detector is illustrated in Figure 2.

An important condition for LRA detectors is that $\boldsymbol{s}$ is taken from a specified complex integer set, i.e., $\boldsymbol{s} \in \mathsf{S}^{\mathcal{M}} \in \mathbb{Z}^{\mathcal{M}}$. And since $\boldsymbol{T}^{-1}$ contains only integers, $\boldsymbol{u} \in \boldsymbol{T}^{-1}\mathsf{S}^{\mathcal{M}}$ is also a complex integer vector. Therefore, after soft estimates $\tilde{\boldsymbol{u}}$ are obtained by LRA linear detectors, a simple suboptimal quantization operation is used to round off the real and imaginary parts of each element $\tilde{u}_n$ in $\tilde{\boldsymbol{u}}$ to the nearest integer as $\hat{u}_n = \lceil \mathfrak{R}\{\tilde{u}_n\} \rfloor + \imath \lceil \mathfrak{I}\{\tilde{u}_n\} \rfloor$, where $\imath = \sqrt{-1}$. The detected signal vector is then obtained by transforming $\hat{\boldsymbol{u}}$ into $\tilde{\boldsymbol{s}}$ by the transform matrix $\boldsymbol{T}$, i.e., $\tilde{\boldsymbol{s}} = \boldsymbol{T}\hat{\boldsymbol{u}}$.

It is worth noting here that when $\boldsymbol{s}$ belongs to the Quadrature Amplitude Modulation (QAM) constellation, then in order to create a complex integer set $\mathsf{S}^m \in \mathbb{Z}^m$, proper shifting and scaling are necessary. Detailed description of these shifting and scaling operations can be found in [7] and references therein.

3.3 Detection Error and Quantization Error

In this section, we analyze the relation between the detection error and quantization error occurred in a LRA linear detector. It is reminded that, in a LRA-detector, soft estimates of the transmitted signal vector are given by the following transformation: $\tilde{\boldsymbol{s}} = \boldsymbol{T}\hat{\boldsymbol{u}}$. The estimated vector $\tilde{\boldsymbol{s}}$ is then quantized to the signal constellation as $\hat{\boldsymbol{s}} = \mathcal{Q}\{\tilde{\boldsymbol{s}}\}$. Due to the effects of additive noise, channel fading as well as rounding operation, errors can occur in this quantization step, making $\hat{\boldsymbol{s}}$ not belong to the original set S^m. In other words, there exist *quantization errors* due to quantization in LRA detectors. To better understand about the quantization error, let us consider a simple example of a 2×2 MIMO system with

$$\boldsymbol{T} = \begin{bmatrix} -\imath & 1-\imath \\ 1 & 1 \end{bmatrix}.$$

If 4-QAM modulation is used then the symbol alphabet after shifting and scaling will be $\mathsf{S} = \{0, 1, \imath, 1+\imath\}$. Let $\boldsymbol{s} = \begin{bmatrix} 1 & 1+\imath \end{bmatrix}^T$ be the transmitted vector, then $\boldsymbol{u} = \boldsymbol{T}^{-1}\boldsymbol{s} = \begin{bmatrix} 1 & \imath \end{bmatrix}^T$. Assume that the quantized vector after detection is $\hat{\boldsymbol{u}} = \begin{bmatrix} 1+\imath & \imath \end{bmatrix}^T$, then the final estimate $\hat{\boldsymbol{s}}$ can be written as $\hat{\boldsymbol{s}} = \mathcal{Q}\{\boldsymbol{T}\hat{\boldsymbol{u}}\} = \begin{bmatrix} 2 & 1+2\imath \end{bmatrix}^T$. Since both the complex integers 2 and $2\imath$ do not belong to the original set S, it is said that a quantization error occurs in the quantization step.

In order to investigate the influence of quantization errors on the overall error performance, we have conducted an experiment based on computer simulation. The MIMO system used in experiment has 4 transmit and 4 receive antennas. Monte-Carlo simulation was run, and the number of iterations that have quantization errors and the number of corresponding bit errors for every 100 detection iterations were counted. Table I shows the average number of iterations in which quantization errors occured and the corresponding percentage of bit errors compared to the total number of bit errors for the three LRA linear detectors: LRA-ZF, LRA-MMSE, and LRA-MMSE based on $\bar{\boldsymbol{H}}$ (LRA-MMSE Ext).

It is interestingly noted from Table I that although quantization errors do not occur very often, they account for most of the bit errors. For example, for the LRA-ZF detector, at SNR = 12dB, there are only 5% of detection iterations containing quantization errors on average, but they account for 90% of the total bit errors. Similarly, for the LRA-MMSE detector, at SNR = 12dB, only 2% of detection iterations are affected by quantization errors on average, but this number corresponds to 78% of the total bit errors. Another important observation can be made from this table is that for the LRA-ZF and LRA-MMSE detector and in the medium and high SNR region, most of detection errors occur in few detection iterations. Only for the LRA-MMSE Ext detector, there are almost no quantization errors. This is because the MMSE criterion is optimized with LRA-MMSE Ext detector. In general, we can see that as SNR increases the quantization error rate becomes small and decreases fast. Based on the above observations, it is

Table I
QUANTIZATION ERROR RATE VS PERCENTAGE OF TOTAL ERRORS OCCURRED IN LRA LINEAR DETECTORS.

SNR[dB]	LRA-ZF		LRA-MMSE		LRA-MMSE Ext.	
	Quantization Error Rate	Percentage of Errors	Quantization Error Rate	Percentage of Errors	Quantization Error Rate	Percentage of Errors
6	47	88	21	66	47	14
8	28	87	12	70	3.3	21
10	13	88	5.8	73	1.56	26
12	5.1	90	2.1	78	0.54	32
14	1.5	92	0.6	84	0.1	38
16	0.36	94	0.13	87	0.018	48

desired to have a nonlinear correction scheme or an optimal detector to recover the bit error for those symbols having quantization errors. One possible approach is to use the reduced-complexity near-ML detector as proposed in [10]. However, the corrector in [10] is not suitable for high-order modulations, such as QAM, due to degraded performance and high computational complexity.

In this paper, we propose to use a sphere detector (SD) as a quantization error corrector to re-detect the quantization erroneously estimated symbols of the LRA linear detectors. Since SD gives exactly the same performance of ML detection (with much lower complexity), most of bit errors could be corrected. Note that the quantization error rate is low, hence we need to use SD in only a few iterations. Therefore the additional complexity will be slightly small. In addition, we also propose a *list quantization* scheme to improve the quantization as well as the BER performance of the LRA detectors. Details on these two schemes are presented next in Sections 4 and 5.

4 QUANTIZATION ERROR CORRECTION USING SPHERE DETECTOR

The idea of using a sphere detector as a quantization error corrector is somehow similar to the reduced-complexity maximum likelihood (ML) detector proposed in [10]. Here the quantization error is like a test of detection quality to find the high probably erroneous detected symbols. Every time the test fails, i.e., quantization error occurs, we apply the SD detector to get the optimal estimate and, thus, to recover the bit errors.

Let us recall here that, for linear LRA detectors, bit errors are most likely to occur in the estimated symbol vector $\hat{s}$ that has one or more elements laying outside the signal constellation. Since SD gives the optimal detection, if we apply sphere detection on this vector, it is expected that most of bit errors would be corrected.

The main idea of sphere detection is briefly explained as below. Unlike ML detection (MLD), SD detection does not search all the constellation points but only those points that lie within a hyper-sphere of radius R around the received signal point [11]. Since the number of points searched is limited, the complexity is reduced as compared to ML detection. For ML detection with

sphere constraint R, we have

$$\hat{s} = \arg \min_{\tilde{s}} \left\{ \|y - Hs\|^2 < R^2 \right\}. \tag{9}$$

Based on QR decomposition $H = QR$ and $Q^H Q = I$, (9) can be rewritten as

$$\hat{s} = \arg \min_{\tilde{s}} \left\{ \|y' - Rs\|^2 < R^2 \right\} \tag{10}$$

$$= \arg \min_{\tilde{s}} \left\{ \sum_{i=1}^{N} \|y_i' - \sum_{j=i}^{N} r_{ij} s_j\|^2 < R^2 \right\}, \tag{11}$$

where $y' = Q^H y$ and r_{ij} is the element at row i and column j of the upper triangular matrix R. The solution of (11) can be obtained recursively starting from layer $i = N$ to $i = 1$. For each layer, constellation value s_i satisfying

$$\left\| \frac{z_i}{r_{ii}} - s_i \right\|^2 < \frac{T_i}{r_{ii}^2} \tag{12}$$

are selected as partial ML candidates for layer i, where

$$z_i = \left(y_i' - \sum_{j=i+1}^{N} r_{ij} s_j \right) \tag{13}$$

and

$$T_i = T_{i+1} - \|z_i - r_{ii} s_i\|^2, \quad \text{with } T_M = R^2. \tag{14}$$

Those satisfied values s_i on each layer can be obtained through the constraint in (12), via direct comparison to the QAM constellation. When a new point is found inside the hyper-sphere (at $i = 1$) the radius is updated with the new minimum Euclidean distance and the algorithm continues the search with a new sphere constraint until there is only one point left. Figure 3 shows an example of one SD searching iteration in different layer in a tree diagram. The dark nodes mean that the value T_i of the previous node already exceeds R^2.

Some advantages thanks to using a sphere detector in combination with LRA linear detectors can be enumerated as below.

i) Initial radius optimization can be avoided: One important issue of sphere detection is to choose a suitable initial radius because the complexity of sphere detectors is very sensitive to this initial radius [12]. In LRA linear detectors, since the estimate is already known from the LRA detection stage, the initial distance can be set equal to the distance from the received signal point to the estimated point.
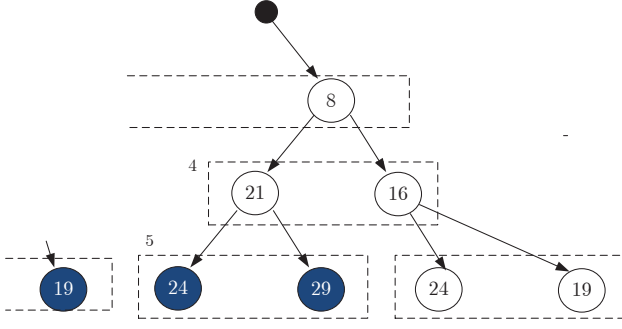
Figure 3. Tree diagram of a sphere detector.

Figure 4. Quantization rate vs limiting number.

ii) Keep the additional complexity small by limit number: In order to keep the additional complexity small, we can limit the number of symbols to be corrected by a limit number. In particular, we use SD for only those estimated symbol $\hat{s}$ whose the number of elements lying outside the signal constellation is smaller or equal to the limit number. This number should be chosen less than $2N$ as we treat the real and imaginary parts separately. In our simulation, we set the limit number equal to 4, 5, and 6 for LRA-ZF, and 3, 4, and 5 for the LRA-MMSE detector. Figure 4 shows the quantization error rates corresponding to different values of the limit number. It is obvious that the quantization error rate is proportional to the number of times that we need to use the sphere detector. This also means that the rate is proportional to the additional complexity. It is clear from the figure that the additional complexity due to sphere detection can be controlled if we use the limit number. For example, if we apply 3 and 5 as the limit numbers for SD-LRA-ZF and for SD-LRA-MMSE, the frequency of using SD is only about 5% for SNR = 8dB.

5 List Quantization LRA Detectors

5.1 Principle of the List Quantization Method

It is known from the previous section that the estimate of the transmit vector is given by

$$\tilde{s} = T\hat{u}. \qquad (15)$$

Thus an error in elements of $\hat{u}$ will also cause an error to $\tilde{s}$ due to nonzero elements of the corresponding column of T. It is interestingly noted that when the error occurs the element-wise differences between the optimum quantization and rounding operation are mainly $+1$, -1, ι or $-\iota$.

Based on that, the idea of the proposed list quantization scheme is explained in brief as follows. Suppose that error occurs in only one component $\hat{u}_i$ of u. Then, the transmitted signal element is mainly expected to be $\hat{u}_i = \hat{u}_i + \text{sign}(\Re\{\tilde{u}_i\} - \lceil\Re\{\tilde{u}_i\}\rfloor)$ or $\hat{u}_i = \hat{u}_i + \iota\,\text{sign}(\Im\{\tilde{u}_i\} - \lceil\Im\{\tilde{u}_i\}\rfloor)$. These candidate elements are the nearest elements of component $\hat{u}_i$ in the real and imaginary axes. If there are N transmit antenna elements, we will have $2N + 1$ candidates of $\hat{u}$, including the original estimated $\hat{u}$. Based on (15), we will have a list of $2N + 1$ candidate symbols $\tilde{s}$. By removing
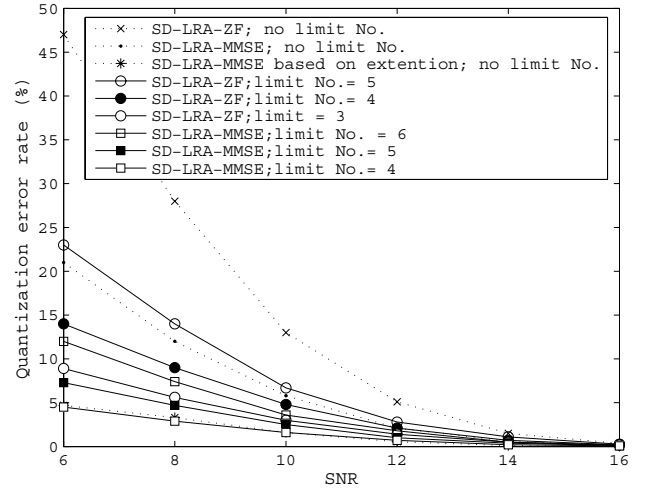
the candidate symbols whose elements lying outside the constellation, we obtain a list of highly reliable candidate symbols. Finally, the symbol nearest to the received point is chosen as the solution of the following equation

$$\hat{s} = \arg\min_{\tilde{s}} \|y - H\tilde{s}\|. \qquad (16)$$

Since LRA detection with the LLL-reduced matrix usually gives a good estimate, we may assume that there is only one erroneous element in $\tilde{u}$. Under this assumption, we can generalize a list-quantization scheme for the LRA detection as follows:

Step 1: Create $2N$ candidate symbols from the estimate of LRA detection. For i from 1 to N, add 2 candidates to the list, changing element $\hat{u}_i$:

$$\hat{u}_i = \hat{u}_i + \text{sign}(\Re\{\tilde{u}_i\} - \lceil\Re\{\tilde{u}_i\}\rfloor)$$

and

$$\hat{u}_i = \hat{u}_i + \iota\,\text{sign}(\Im\{\tilde{u}_i\} - \lceil\Im\{\tilde{u}_i\}\rfloor).$$

Step 2: Remove those candidates lying outside the signal constellation to achieve the list of candidates. If the list is empty, quit and return the estimate of LRA detection as the final result.

Step 3: Calculate the Euclidean distance from the received point to the LRA-based estimate point, update the distance for candidates in the list.

Step 4: Decide the best candidate which has the shortest distance.

5.2 Complexity Analysis

In order to keep the additional complexity small, we take a further look into the main step of the scheme, i.e., the Euclidian distance calculation. In most cases, the candidate symbols from the list are different from the first LRA estimate symbols by only one or two elements. Therefore, to calculate the new distance, we only need to calculate the different part of the term $(y - H\hat{s})$ inside the $\|.\|$. For example, if $\hat{s}_1$ and $\hat{s}$ are

different in the i-th element, the distance from $\hat{s}_1$ can be calculated as:

$$D_1 = \|y_i - \{h_1\hat{s}_1(1) + ... + h_N\hat{s}_1(N) + h_i\hat{s}_1(i)\}\| \quad (17)$$
$$= \|\{y_i - h_1\hat{s}(1) - ... - h_N\hat{s}(N)\} - h_i\hat{s}_1(i)\| \quad (18)$$

where $\hat{s}_1(i)$ represents the ith element of $\hat{s}_1$. We can see that the term inside the bracket of the second equation is already known when calculating the distance for $\hat{s}$. The only calculation needed is the multiplication and addition operations when calculating $h_i\hat{s}_1(i)$ and the norm. Moreover, we can represent the distance in the real domain using new variables

$$\bar{H} = \begin{bmatrix} \Re\{H\} & -\Im\{H\} \\ \Im\{H\} & \Re\{H\} \end{bmatrix} \quad (19)$$

and the real-valued vectors

$$\bar{y} = \begin{bmatrix} \Re\{y\} \\ \Im\{y\} \end{bmatrix}, \quad \bar{s} = \begin{bmatrix} \Re\{s\} \\ \Im\{s\} \end{bmatrix}. \quad (20)$$

The real matrix $\bar{H}'$ has the size $m \times n$ with $m = 2M$ and $n = 2N$. The real vectors $\bar{y}$ and $\bar{s}$ have size of $2N \times 1$. Then the distance can be expressed in the real domain as follows

$$D_1 = \|\{\bar{y}_i - \bar{h}'_1\bar{\hat{s}}(1) - ... - \bar{h}'_{2N}\bar{\hat{s}}(2N)\} - \bar{h}'_i\bar{\hat{s}}_1(i)\|.$$

It is clear that when comparing with $h_i\hat{s}_1(i)$, the number of required multiplication and addition operations in $\bar{h}'_i\bar{\hat{s}}_1(i)$ is reduced by half. For calculating the distance for signal point $\hat{s}$, we need $(2N+1)2M$ multiplications operations. However, to update the distance D_1 for the point $\hat{s}_1$, we need only $4M$ multiplication operations. This means that if the expected number of different elements in one candidate symbol is $\bar{n}_{\text{dif}}$ then in order to update the distance we need $4\bar{n}_{\text{dif}}M$ multiplication operations. If the expected number of candidate symbols in the list is $\bar{N}_{\text{list}}$, we need $4\bar{N}_{\text{list}}\bar{n}_{\text{dif}}M$ multiplications, where $\bar{n}_{\text{dif}}$ is the number of different elements in average. Because step 2 makes the numbers $\bar{N}_{\text{list}}$ and $\bar{n}_{\text{dif}}$ become quite small, the complexity is negligible. For example, with 4×4 MIMO, at SNR= 10dB, the number value is around 2 for LRA-ZF detection and 3 or 4 for LRA-MMSE detection (see Figure 10).

6 Simulation Results

In order to evaluate performance of the proposed schemes, we have carried out simulations for a 4×4 MIMO system. The channel was assumed to be the flat Rayleigh fading with its complex gain generated using complex Gaussian random variables with zero mean and unit variance.

6.1 Results for SD-LRA

BERs of the LRA-ZF/MMSE and SD-LRA-ZF/MMSE detectors are shown in Figures 5 and 6 for different limiting numbers in a 4×4 MIMO system. We use 4 and 5 for ZF, and 5 and 6 for MMSE, in this simulation. In these figures, we compare BERs obtained using conventional LRA and combined SD-LRA based on
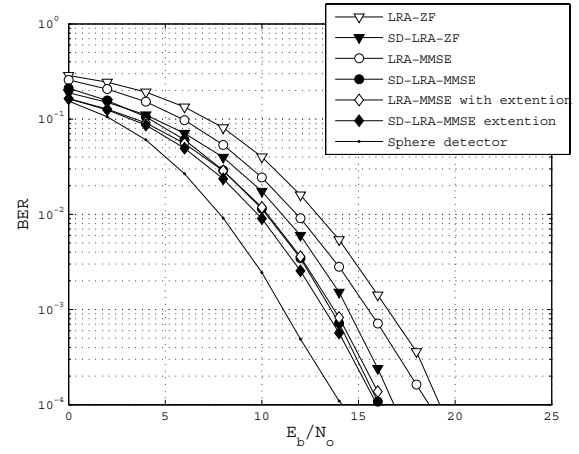

Figure 5.   BER performance of LRA-ZF/MMSE and SD-LRA ZF/MMSE detection, MIMO 4×4, limit number for ZF = 5, limit number for MMSE = 6.
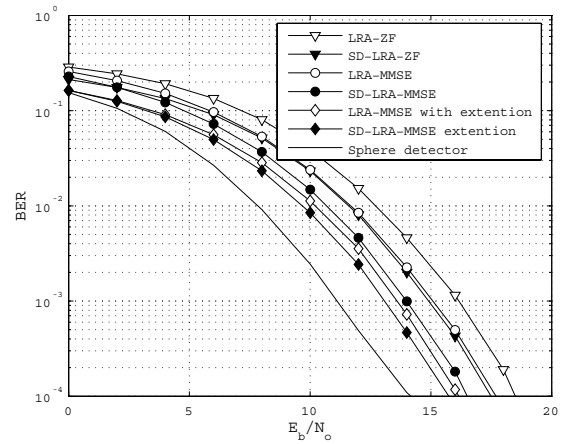

Figure 6.   BER performance of LRA-ZF/MMSE and SD-LRA ZF/MMSE detection, MIMO 4×4, limit number for ZF = 4, limit number for MMSE = 5.

ZF and MMSE. It can be seen from the figures that using the quantization error correction scheme, the BER performance of the LRA-ZF and LRA-MMSE detectors are improved steadily. The LRA-MMSE based on $\bar{H}'$ does not benefit from using SD because there are almost no quantization errors occured as we have already discussed from the previous section. It is expected that the additional complexity for SD-LRA-ZF and SD-LRA-MMSE goes to nearly zero in the high SNR region.

6.2 Results for List Quantization LRA Detectors

BERs of the LRA-ZF/MMSE detectors with and without list quantization are compared in Figure 7 for a 4×4 MIMO system. It can be seen from the figure that, when the list quantization scheme is used, the BER performance of the LRA-ZF/MMSE detector is improved significantly. At BER = 10^{-3}, the list quantization LRA-ZF detector can gain nearly 2.5 dB over the conventional LRA-ZF. The performance gain of LRA-MMSE based on $\bar{H}$ is about 1.5 dB. It is clear that the LRA-MMSE based on $\bar{H}$ with list quantization can achieve near-ML BER performance. We compare the list-quantization method for two cases in Figure 8. First, a list of candidates
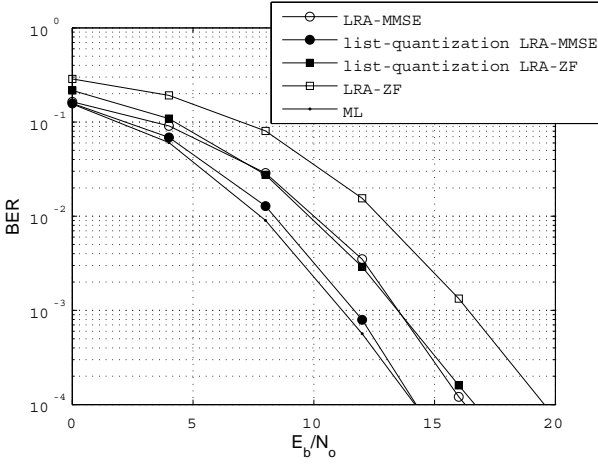
Figure 7.   BER performance of LRA-ZF/MMSE detection with and without list-quantization method, MIMO 4×4, QAM modulation.
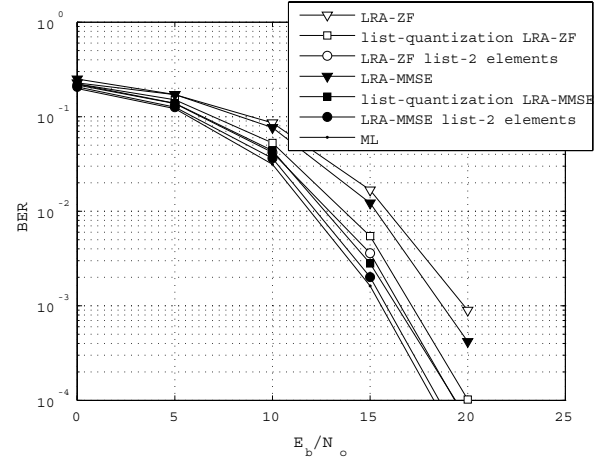

Figure 9.   BER performance of LRA-ZF/MMSE detection with and without list-quantization method, MIMO 4×4, 16QAM modulation, list of candidates is the combination of 1 or 2 elements.
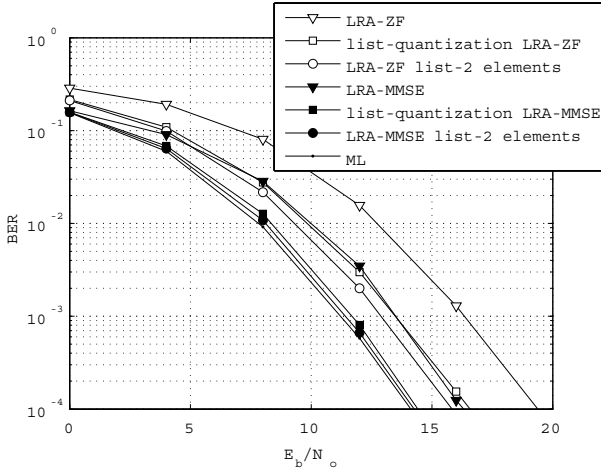

Figure 8.   BER performance with and without list-quantization method, MIMO 4×4, 4QAM modulation, list of candidates is the combination of 1 or 2 elements.
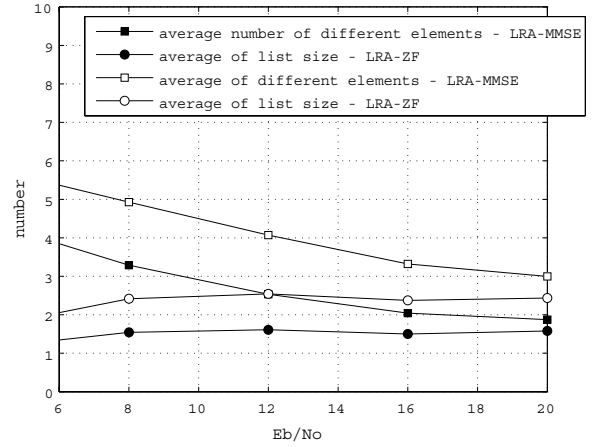

Figure 10.   The average size of the list and average number of different elements, MIMO 4×4, QAM modulation.

is generated by changing only one element of the estimate. Second, the list of candidates is generated by changing one and two elements of the estimate. The improvement is almost the same in both cases. Therefore, it can be concluded that the list quantization scheme will work efficiently enough by changing only one element. That also supports the fact that in most cases the error appears in one element in the estimate $\hat{u}$. From Figure 9, the impact of list quantization scheme is even more clear with 16QAM modulation. The LRA-ZF detector with list quantization even outperforms the conventional LRA-MMSE. Similarly, in both cases of changing 1 or 2 elements, we get almost the same performance.

Finally, Figure 10 shows the averaged number of candidate symbols in the list and the number of different elements. Based on the complexity analysis from the previous section qand the value in the figure, the number of multiplication and addition operations is just several multiples of number of transmit antenna elements. In other words, it is comparable with the complexity of calculation of the Euclidian distance for one signal point. We can see that the average values

of the n_{dif}, N_{list} are small, especially in the high SNR region. Therefore, if the number of antenna elements is not so large, the additional complexity is small and negligible. With the above results, it is reasonable to conclude that the performance is significantly improved for both LRA-ZF and LRA-MMSE detectors with very low additional complexity.

7 Conclusion

In this paper we propose two methods to improve the quantization step in LRA linear detectors. Firstly, we have investigated the relation between detection error and the quantization error of the LRA-ZF and LRA-MMSE detectors. We have shown that there is a strong relation between error in quantization and detection error. Based on this observation, in order to reduce the number of bit errors, we apply SD for high probably erroneous estimate to achieve its optimal solution. The SD-LRA detectors show the improvement over the LRA linear detectors. By using a complexity limit number, we can keep the complexity reasonable small, especially at the high SNR.

Secondly, we also proposed the list-quantization method for LRA-linear detection. We have shown that list-quantization can limit the error in quantization step and, thus, reduce the detection error. Simulation results showed that a significant improvement can be achieved for both LRA-ZF and LRA-MMSE detectors. Especially, the LRA-MMSE based on $\bar{H}$ with list quantization can get near-ML BER performance. The proposed quantization method will be more efficient with a higher constellation size. The simulation result shows that the method is efficient enough to generate the list of candidates by changing only one element of the estimate. We also calculate the complexity needed for list-quantization scheme. We showed that the additional complexity is just several multiples of number of transmit antenna elements. Thus, additional complexity will be negligible, especially in high SNR range.

Acknowledgement

References

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[2] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture," *Electronics Letters*, vol. 35, no. 1, pp. 14–16, Jan. 1999.

[3] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 1, Cambridge, MA, USA, 17-21 Nov. 2002, pp. 424–428.

[4] D. Wübben, R. Böhnke, V. Kühn, and K.-D. Kammeyer, "Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction," in *Proc. IEEE Int. Conference on Communications (ICC)*, vol. 2, 20-24 June 2004, pp. 798–802.

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

[6] K. L. Clarkson, W. Sweldens, and A. Zheng, "Fast multiple-antenna differential decoding," *IEEE Transactions on Communications*, vol. 49, no. 2, pp. 253–261, Feb. 2001.

[7] T. D. Nguyen, X. N. Tran, and T. Fujino, "Layer error characteristics of lattice-reduction aided V-BLAST detectors," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 10, pp. 2535–2543, Oct 2006.

[8] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. IEEE Int. Conference on Acoustic, Speech and Signal Processing*, Istanbul, Turkey, 2000, pp. 5–9.

[9] C. Windpassinger and R. F. H. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proc. IEEE Information Theory Workshop (ITW)*, Erlangen, Germany, Mar. 31 - Ap. 4 2003, pp. 345–348.

[10] A. T. Le, X. N. Tran, and T. Fujino, "Combined ML-MMSE multiuser detection for STBC-OFDM," *IEICE Transactions on Fundamentals*, vol. E88-A, no. 10, 2005.

[11] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, Jul 1999.

[12] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, pp. 389–399, Nov 2003.
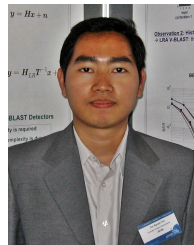
Tien Duc Nguyen received his bachelor of engineering degree in electronics and telecommunication from Hanoi University of Technology, Vietnam, in 2001, master of engineering in Mobile communications from Aalborg University, Denmark in 2004. During the period from 2005 to 2008 he was a doctorate student at The University of Electro-Communications, Japan. Mr. Nguyen passed away in 2009 due to an accident during finalizing this work.

Tadashi Fujino was born in Osaka, Japan on 15 July, 1945. He received B.E. and M.E. degrees in electrical engineering and Dr. Eng. degree in communication engineering from Osaka University, Osaka, Japan, in 1968, 1970, and 1985, respectively. Since April 2011, he has been Professor Emeritus of The University of Electro-Communications (UEC), Tokyo, Japan. From April 2000 to March 2011, he was an ordinary professor in wireless communications at Department of Information and Communication Engineering and Informatics, Graduate School of Informatics and Engineering, The UEC, Tokyo, Japan. Before then, he had been with Mitsubishi Electric Corporation, Tokyo, Japan since 1970, where he devoted in R&D in the wireless communications area such as digital satellite communications and digital land mobile communications. His major works include the feasibility study and the hardware development of the 120 Mbps trellis coded (TC) 8-PSK burst modem for TDMA system. This work was done at INTELSAT's request, and is the first development in the world. In the above feasibility study and the hardware development, he made two innovative inventions: one is the "pipe-lined Viterbi decoder," which makes possible for TC modems to operate in very high bit rates of hundreds Mbps, and the other is the clock aided carrier recovery system for the coherent detection of M-PSK modulated signal, which generates very little pattern noise. After designing and manufacturing the hardware of the TC burst modem using these inventions, he made field test, which was successfully done. By his two inventions and the field test results, it is verified that the high bit rate TC burst modem should be feasible to design. His current interests include the signal detection in MIMO systems such as lattice-reduction aided detection. He wrote a single authored book "Digital mobile communication," and three co-authored books. He received Meritorious Award from the ARIB (The Associate of Radio Industries and Businesses of Japan) of MPT of Japan, in 1997. He is also a Visiting Professor at Bac-Ha International University, Hanoi, Vietnam. Dr. Fujino is a Fellow of IEEE and a member of IEICE.

Xuan Nam Tran is currently an associate professor at Department of Communications Engineering at Le Quy Don Technical University Vietnam. He received his master of engineering (ME) in telecommunications engineering from University of Technology Sydney, Australia in 1998, and doctor of engineering in electronic engineering from The University of Electro-Communications, Japan in 2003. From November 2003 to March 2006 he was a research associate at the Information and Communication Systems Group, Department of Information and Communication Engineering, The University of Electro-Communications, Tokyo, Japan.

Dr. Tran research interests are in the areas of adaptive antennas, space-time processing, space-time coding and MIMO systems. He is the author and co-author of more than thirty technical papers published in international journals and conference proceedings. Dr. Tran is a recipient of the 2003 IEEE AP-S Japan Chapter Young Engineer Award. He is a member of IEEE, IEICE, and the Radio-Electronics Association of Vietnam.