2006 IEEE International
Symposium on Signal Processing
and Information Technology

A Sphere Decoding Algorithm for MIMO Channels

Chin-Yun Hung and Tzu-Hsien Sang
Department of Electronics Engineering, National Chiao Tung University
1001, Ta Hsueh Road, Hsinchu 300, Taiwan, R.O.C.,
ferra.ee93g@nctu.edu.tw, tzuhsien54120@faculty.nctu.edu.tw

Submission topic: Signal Processing for Communications

***Abstract*— Multi-Input Multi-Output (MIMO) transmission has become a popular technique to increase spectral efficiency. Meanwhile, the design of cost-effective receivers for MIMO channels remains a challenging task. Maximum-Likelihood (ML) detector can achieve superb performance, yet the computational complexity is enormously high. Receivers based on sphere decoding (SD) reach the performance of ML detectors, and potentially a great deal of computational cost can be saved. In this paper, a practical sphere-decoding algorithm is proposed. It utilizes a simple and effective way to set the initial radius which plays a decisive role in determining the computational complexity. Furthermore, a pseudo-antenna augmentation scheme is employed such that SD can be applied where the number of receive antennas is less than that of transmit antennas; thus enhance the applicability of this powerful algorithm.**

I. Introduction

Sphere Decoding has recently been applied to signal detection problem for Multiple-Input Multiple-Output (MIMO) systems [1]–[4]. It is a reduced search algorithm for doing ML detection. Notice that brute-force ML detection has computational complexity that is exponentially growing in the number of sub-streams, the constellation size, and the number of transmit antennas; as a result, it is not feasible for practical systems. Indeed, SD holds the potential of significantly reducing the computational cost while maintaining the superb performance of an ML detector and therefore is compared favorably with other sub-optimal detectors proposed for MIMO systems.

Some noteworthy detection algorithms for MIMO systems are reviewed in the following.

1) Linear detection methods: The linear detection method first estimates the channel matrix then tries to compensate (inverse) the channel effect by another matrix. The inverse matrix is usually based on Zero Forcing (ZF) or Minimum Mean Square Error (MMSE) criterion. This method requires very low computational complexity, but results in significant performance degradation.

2) Successive Interference Cancellation (SIC): Successive interference cancellation peels the transmission signal apart one data stream at a time. It decodes and cancels the data stream iteratively until all transmitted streams are resolved. If sorting is done to determine the decoding order from the highest to the lowest SNR, it is called ordered successive interference cancellation (OSIC). An example is the so-called Vertical Bell-laboratory LAyered Space-Time (V-BLAST) receiver [5]. OSIC has a slightly better performance than SIC does, but is still suboptimal and suffers from error propagation.

3) Brute-Force Maximum Likelihood(ML) Detection: Assuming that the transmitted data sequence is i.i.d., the maximum likelihood detector for a MIMO system performs the operation:

$$\begin{aligned}\hat{\mathbf{s}}_{ML} &= \underset{\mathbf{s}\in\mathcal{O}^{N_T}}{\arg\min} \left\|\mathbf{y}-\mathbf{H}\mathbf{s}\right\|^2 \\ &= \underset{\mathbf{s}\in\mathcal{O}^{N_T}}{\arg\min} (\mathbf{y}-\mathbf{H}\mathbf{s})^H(\mathbf{y}-\mathbf{H}\mathbf{s})\end{aligned} \quad (1)$$

where $\mathbf{y}$ is the observed vector signal, $\mathbf{H}$ is the channel matrix whose size is $N_R \times N_T$, $\mathcal{O}^{N_T}$ is the entire set of possible transmitted vector symbols, $\mathcal{O}$ is the complex-valued modulating constellation, and $(\cdot)^H$ means Hermitian transpose. The ML detector is optimal in terms of symbol error rate, but the computational complexity can be prohibitively high if it is implemented by exhaustively searching over $\mathcal{O}^{N_T}$.

4) Sphere Decoding: In 1985, U. Fincke and M. Phost proposed an algorithm named Fincke-Phost algorithm [6] (or SD algorithm) which offers a large reduction in computational complexity for the class of computationally-hard combinatorial problems, for instance, the aforementioned ML detection problem. SD algorithm used for resolving MIMO channel was presented in [1]–[4] and was shown to reduce the complexity of ML detector significantly [1], [7], [8]. The enormous computational complexity of ML detector arises from the huge number of vector symbols to be compared in order to find the solution in (1). The main idea of SD algorithm is to use a highly efficient method to reduce the number of candidate vector symbols before the actual comparison happens. For more on the efficiency of SD, please refer to [1], [7], [8].

Let $\mathcal{D}$ be a sphere centered at the received vector $\mathbf{y}$, and the radius d of $\mathcal{D}$ is properly defined such that only a small number of vector symbols fall inside $\mathcal{D}$ after being transformed by the channel matrix. The search of the closest transformed vector symbol to $\mathbf{y}$ can be conducted among these candidates in $\mathcal{D}$ rather than the entire set $\mathcal{O}^{N_T}$. A well-designed sphere decoder would have performance equal to that of an ML detector. For example, it can reach full diversity while V-BLAST can only reach $N_R - N_T + 1$ [9].

Two questions need to be addressed for an effective sphere decoder to be constructed:

1. How to choose the radius d such that the number of candidates is well limited?

2. How to determine efficiently if a channel symbol actually lies inside the hypersphere $\mathcal{D}$?

In this paper, a simple yet effective method to set the radius of the hypersphere $\mathcal{D}$ is proposed. A pseudo-antenna augmentation scheme is also proposed such that SD can efficiently determine the position of a lattice point relative to $\mathcal{D}$ in the case where the number of transmit antennas is larger than the number of receive antennas, thus expand the applicability of SD. Compared to existing literatures which handle rank-deficient channel matrices [10]–[12], our method is more intuitive and straightforward, and it enjoys a computational complexity in polynomial when SNR is sufficiently high, while methods in [10], [11] have a complexity growing exponentially in $(N_T - N_R)$.

The rest of this paper is organized as follows. The MIMO system model and basic formulations are laid out in Section II. The radius-setting method and the pseudo-antenna augmentation scheme are described in Section III. Simulation results are presented in Section IV, and finally, a brief conclusion in Section V.

II. System Model and Basic SD

The MIMO system model is as follow. Assume N_T transmit antennas and N_R receive antennas, the received signal is given by

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n} \tag{2}$$

where $\mathbf{y} \in \mathcal{C}^{N_R}$ is the received signal vector, $\mathbf{H} \in \mathcal{C}^{N_R \times N_T}$ is the Rayleigh flat fading channel matrix, $\mathbf{s}$ is the transmitted vector symbol in $\mathcal{R}^{N_T}$ or $\mathcal{C}^{N_T}$ depending on the modulation scheme, and the entries of $\mathbf{n}$ is the additive i.i.d. zero mean circularly symmetric complex Gaussian (ZMCSCG) noise with variance of σ^2, i.e., $n_k \sim CN(0, \sigma^2), k = 1, ..., N_R$. The lattice point $\mathbf{Hs}$ lies inside the hypersphere $\mathcal{D}$ of radius d if and only if

$$d^2 \geq \left\|\mathbf{y} - \mathbf{Hs}\right\|^2 . \tag{3}$$

Assume that $\mathbf{H}$ is column-independent (i.e. $N_R \geq N_T$) and $\mathbf{H}$, $\mathbf{y}$, $\mathbf{s}$ and $\mathbf{n}$ are real-valued, then $\mathbf{H}$ can be QR-factorized [13] as

$$\begin{aligned}\mathbf{H} &= \mathbf{QR} \\ &= \begin{bmatrix}\mathbf{Q_1} & \mathbf{Q_2}\end{bmatrix}\begin{bmatrix}\mathbf{R}' \\ \mathbf{0}\end{bmatrix}\end{aligned} \tag{4}$$

where $\mathbf{Q} \in \mathcal{R}^{N_R \times N_R}$ is an orthonormal matrix, $\mathbf{R} \in \mathcal{R}^{N_R \times N_T}$ is an upper triangular matrix, and $\mathbf{R}'$ is an $N_T \times N_T$ upper triangular matrix of $\mathbf{R}$. The matrices $\mathbf{Q_1}$ and $\mathbf{Q_2}$ consist of the first N_T and last $N_R - N_T$ orthonormal columns of $\mathbf{Q}$ respectively. ¿From (3) and (4), we have

$$d^2 - \left\|\mathbf{Q_2}^H\mathbf{y}\right\|^2 \geq \left\|\mathbf{Q_1}^H - \mathbf{Rs}\right\|^2 . \tag{5}$$

Define $d'^2 = d^2 - \|\mathbf{Q_2}^H\mathbf{y}\|^2$ and $\mathbf{z} = \mathbf{Q_1}^H\mathbf{y}$, and (5) becomes

$$d'^2 \geq \sum_{i=1}^{N_T}\left(z_i - \sum_{j=i}^{N_T} r_{i,j}s_j\right)^2 , \tag{6}$$

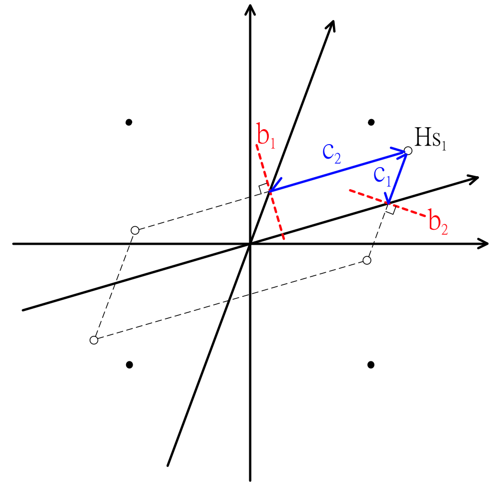

Fig. 1. The diagram shows the idea of finding a proper radius. Assume BPSK and a 2×2 channel matrix for simplicity.

which is the primary working equation in SD to decide whether a lattice point falls inside $\mathcal{D}$ and hence is qualified as a candidate. Next, the set of all candidates is searched and the one closed to the received signal vector is chosen to generate the decoding result [1], [3], [9], [14].

If $\mathbf{H}$, $\mathbf{y}$, $\mathbf{s}$, and $\mathbf{n}$ are complex-valued, they can be written as

$$\begin{aligned}\mathcal{H} &= \begin{bmatrix}\Re\{\mathbf{H}\} & -\Im\{\mathbf{H}\} \\ \Im\{\mathbf{H}\} & \Re\{\mathbf{H}\}\end{bmatrix} \\ \mathcal{Y} &= \begin{bmatrix}\Re\{\mathbf{y}\} \\ \Im\{\mathbf{y}\}\end{bmatrix} \\ \mathcal{S} &= \begin{bmatrix}\Re\{\mathbf{s}\} \\ \Im\{\mathbf{s}\}\end{bmatrix}\end{aligned} \tag{7}$$

where $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represents the real part and image part respectively. Then we can use (7) in substitution for $\mathbf{H}$, $\mathbf{y}$, and $\mathbf{s}$ in (3) and (4). Other modification schemes exist for complex values in certain modulation schemes [4], [15]; our proposed SD algorithm should work well with these alternatives.

III. The Proposed SD Algorithm

To make the SD algorithm a practical choice for MIMO receiver design, two important modifications are proposed. The first is for finding a proper value for d and the second is a pseudo-antenna augmentation scheme to expand the applicable scope of SD. These modifications are discussed in the following two subsections.

A. Setting the Radius d

In drawing the decision regions for an ML detector, the decision boundaries lie on the mid-lines between neighboring lattice points. If the shortest decision distance is used as the initial value of d, it is most likely that the SD algorithm finds one and only one candidate in the hypersphere $\mathcal{D}$ when the noise is small enough that no decision errors occur (this is the case for most of the time).

The shortest decision distance can be easily calculated for certain highly regular modulation constellations. For instance, the shortest decision distance in a square lattice is

$$\min_{i \neq j} \frac{1}{2} \left\| (\mathbf{H}\mathbf{s}_i - \mathbf{H}\mathbf{s}_j) \right\| \tag{8}$$

where $\mathbf{s}_i$ and $\mathbf{s}_j \in \mathcal{O}^{N_T}$ are the transmitted symbol vectors. For square QAM, the minimum decision distance can be found as

$$\min_{k_i, l_i} \frac{1}{2} \left\| \sum_{i=1}^{N_T} \left[(-1)^{k_i} - (-1)^{l_i} \right] (\mathbf{H})_i \right\| \tag{9}$$

where $(\mathbf{H})_i$ denotes the i-th column of $\mathbf{H}$, k_i and l_i takes on the value 0 or 1, and the vector $[k_1, ..., k_{N_T}] \neq [l_1, ..., l_{N_T}]$. The expression of minimum distance can be further simplified as

$$\min_{k \in (1, \cdots, 3^{N_T} - 1)} \left\| \sum_{i=1}^{N_T} c_{k\_i} (\mathbf{H})_i \right\| \tag{10}$$

where $[c_{k\_1}, \cdots, c_{k\_N_T}]$ represents all possible non-zero vectors whose elements take on values from $\{0, 1, -1\}$. Therefore, to find the minimum decision distance is to find the minimum norm over a set of random vectors with complex Gaussian elements.

To find the minimum norm in (10) is straightforward; nevertheless, it can take a long time if the problem dimension is large. Notice that among these random vectors, $(H)_1, \cdots, (H)_{N_T}$ have the smallest expected norm. As a result, when N_T is large, the minimum norm will likely occur as the norm of some vector in $\{(H)_1, \cdots, (H)_{N_T}\}$. Therefore, it is proposed that, instead of the minimum decision distance, the minimum column norm in (11) is used as the initial value of d. If no candidate point is found inside the hypersphere, then a larger value will be adopted and the SD procedure is repeated until a termination criterion is met. In short, we make

$$d_{initial} = \min_i \|(\mathbf{H})_i\|, \quad i = 1, ..., N_T. \tag{11}$$

Fig. 1 shows the concept with a simple example of a 2×2 $\mathbf{H}$. Solid points represent the transmitted QPSK symbols, and circles are the received lattice points, i.e., the transmitted symbols multiplied by the channel matrix. Line b_1 and b_2 represent the mid-lines between neighboring points, and c_1 and c_2 are the two decision distances of $\mathbf{H}\mathbf{s_1}$. In this example, c_1 and c_2 are exactly the column norms of $\mathbf{H}$, and c_1 is chosen as the initial radius of hypersphere $\mathcal{D}$.

B. A Pseudo-Antenna Augmentation Scheme

Typical sphere decoders for MIMO channels can only handle the case where $N_R \geq N_T$ [1]. These sphere decoders fail when $N_T > N_R$ because $\mathbf{H}$ does not have full column rank and therefore cannot be QR-factorized. Here, a modification is proposed to deal with the case $N_T > N_R$.

The idea is to augment $\mathbf{H}$ into a matrix with full column rank. Let the augmented matrix be

$$\tilde{\mathbf{H}}_{N_T \times N_T} = \begin{bmatrix} a\mathbf{I}_{(N_T - N_R)} & \mathbf{0}_{(N_T - N_R) \times N_R} \\ \multicolumn{2}{c}{\mathbf{H}} \end{bmatrix} \tag{12}$$

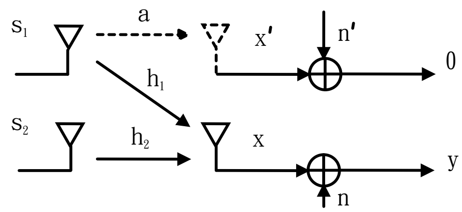

Fig. 2. The diagram of an augmented 2×2 MIMO system.

in which the bottom N_R rows comprise the original channel matrix, $\mathbf{I}$ is the identity matrix, and a is either a small real or complex number depending on the modulation scheme. The pseudo received vector is defined as

$$\begin{bmatrix} as_1 \\ \vdots \\ as_{N_T - N_R} \\ \sum_{i=1}^{N_T} h_{1i} s_i + n_1 \\ \vdots \\ \sum_{i=1}^{N_T} h_{N_R,i} s_i + n_{N_R} \end{bmatrix} \tag{13}$$

and the noise vector is augmented as

$$\begin{aligned} \tilde{\mathbf{n}}_{N_T \times 1} &= \begin{bmatrix} -as_1 \\ \vdots \\ -as_{N_T - N_R} \\ n_1 \\ \vdots \\ n_{N_R} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{n'}_{(N_T - N_R) \times 1} \\ \mathbf{n}_{N_R \times 1} \end{bmatrix} \end{aligned} \tag{14}$$

to make the final augmented received vector to be

$$\begin{aligned} \tilde{\mathbf{y}}_{N_T \times 1} &= \begin{bmatrix} \mathbf{0}_{(N_T - N_R) \times 1} \\ \mathbf{y}_{N_R \times 1} \end{bmatrix} \\ &= \tilde{\mathbf{H}}\mathbf{s} + \tilde{\mathbf{n}}. \end{aligned} \tag{15}$$

By this augmentation, $\tilde{\mathbf{H}}$ has full column rank and can be decomposed via standard QR factorization algorithms. The SD algorithm can now be applied with similar effectiveness for the case $N_T > N_R$. This method is similar but more straightforward than the method in [12] in which an augmented diagonal matrix $\alpha\mathbf{I}$ is added to the matrix $\mathbf{H}^H\mathbf{H}$ to make it full-rank. More comparisons will be made when the effect of a is analyzed.

The concept of pseudo-antenna augmentation is shown in Fig. 2 where a simple 2×1 MIMO channel is augmented to a 2×2 MIMO channel. Fig. 3 shows the space diagram of the transmitted symbol vectors (top), the pseudo received signals (middle), and the augmented received signals (bottom). From (13) and (15), the smaller the value of a is , the closer

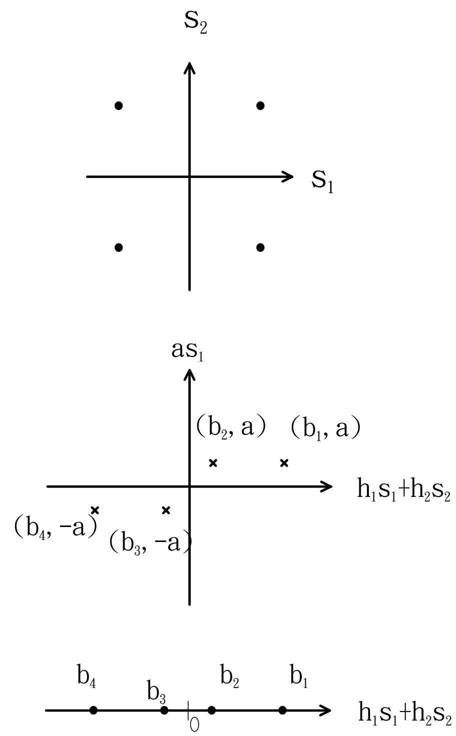

Fig. 3. The space diagram of the transmitted symbol vectors (top), the pseudo received signal vectors (middle), and the augmented received signal vectors (bottom) with the 2×1 MIMO channel. Assume BPSK and $h_1 > h_2 > 0$ for simplicity. Define $b_1 = h_1 + h_2$, $b_2 = h_1 - h_2$, $b_3 = -h_1 + h_2$, $b_4 = -h_1 - h_2$ for convenience.

the augmented and pseudo received signals become. This observation is also shown in Fig. 3.

The effect of the value taken by a can be further analyzed as follows. The set of constellation points resulting in received signals inside the hypersphere $\mathcal{D}$ is found as

$$\mathbf{s}^{\mathcal{D}} = \left\{ \mathbf{x} \middle| d^2 \geq \left\| \tilde{\mathbf{y}} - \tilde{\mathbf{H}}\mathbf{x} \right\|^2 \right\}. \quad (16)$$

The inequality in (16) can be expanded to

$$d^2 \geq |a|^2 \sum_{i=1}^{N_T - N_R} |s_i|^2 + \sum_{i=1}^{N_R} \left| \sum_{j=1}^{N_T} h_{ij}(s_j - x_j) + n_i \right|^2. \quad (17)$$

The lower bound of the radius d with which the correct symbol $\mathbf{s}$ lies in the hypersphere, i.e., $\mathbf{x} = \mathbf{s} \in \mathbf{s}^{\mathcal{D}}$, depends on the noise condition and a. Assume QPSK for simplicity, then $|s_1|^2 = ... = |s_{N_T - N_R}|^2 = 2$ and the lower bound in (17) satisfies

$$d_{LB}^2 = 2(N_T - N_R)|a|^2 + \sum_{i=1}^{N_R} |n_i|^2. \quad (18)$$

The expected lower bound is thus

$$E\{d_{LB}^2\} = 2(N_T - N_R)|a|^2 + N_R \sigma^2. \quad (19)$$

As can be seen clearly in (18), if a is small, the lower bound on the radius with which the correct symbol vector can be included is essentially independent of a. But if a is large, the radius needs to be large.

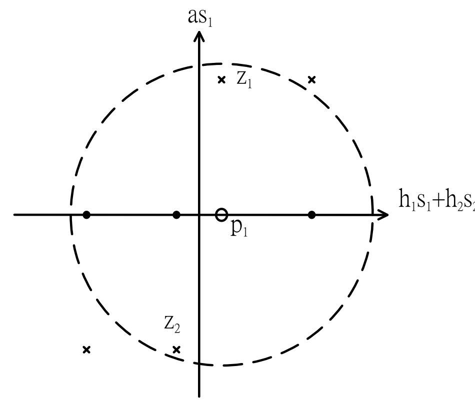

Fig. 4. The space diagram of the hypershpere $\mathcal{D}$ when a is very large. Assume BPSK and a 2×1 MIMO channel for simplicity.

Fig. 4 shows the diagram of a simple example with a 2×1 MIMO channel, BPSK, and a large a. Let point p_1 be the augmented received signal and z_1 the pseudo received signal. The total number of possible received points is 4. As is said before, the radius of the sphere needs to be large. However, when setting the radius, it is extremely difficult for the decoder to find a radius barely large enough to include the lattice point corresponding to the correct symbol while avoid including wrong lattice points in the sphere simultaneously. In Fig. 4, the sphere not only contains the correct point z_1 but also z_2. If a more sophisticated modulation such as 64-QAM is used, and the number of transmit antenna is larger, much more lattice points will inevitably be included in the large hypershpere, and the efficiency of SD will be greatly diminished. Therefore, a should be as small as possible, as long as the numerical stability is maintained in the computing process. With a small a, the complexity of SD is essentially independent of a and the same as that of usual SD algorithms, i.e., roughly $O({N_T}^3)$ when SNR is high [1]. The efficiency of the method in [12], on the contrary, depends on the choice of α, and the optimal choice of α depends on noise condition and is not easy to find.

IV. Simulation results

Fig. 5 shows the performance of SD compared to that of ML receiver. The value of a is set to be very small and the BER performance is equal to that of a brute-force ML receiver.

Fig. 6 shows the average number of candidates found in $\mathcal{D}$ when different values of a and $\frac{E_b}{N_0}$ and the proposed initial radius are used. Notice that when a is getting smaller, say, less than $0.1 + 0.1j$, the number of candidates found in $\mathcal{D}$ is essentially independent of a and is only function of SNR. Also notice that when SNR is moderately large, e.g., in the applications of spatial multiplexing, the number of candidates is close to 1. This means the proposed SD algorithm is operating in a very efficient manner.

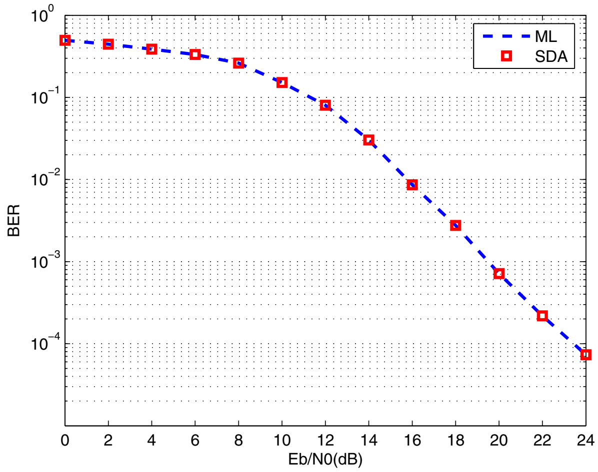

Fig. 5. The BER curves of SD and brute-force ML detector. Assume $N_T = 6$, $N_R = 3$, QPSK, spatial multiplexing, and $a = 0.1 + 0.1j$.

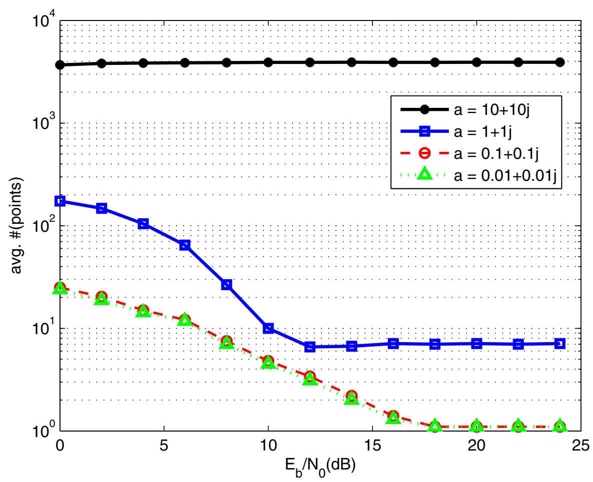

Fig. 6. The average number of candidates inside sphere D with different ralues of value of a and $\frac{E_b}{N_0}$. Assume $N_T = 6$, $N_R = 3$ and QPSK modulation.

V. Conclusion

SD algorithm can significantly lower the computational cost of ML detectors by reducing the number of possible candidates before executing the final step of exhaustive search. In this paper, two special features are introduced to enhance the capability of SD. First, a radius-setting method is used to keep the number of candidate lattice points consistently low. Second, a pseudo-antenna augmentation scheme is employed to cope with the situation where the number of transmit antennas is large than that of receive antennas, which happens often in real-world applications. In short, the modified SD algorithm constitutes an attractive option for practical MIMO receiver design.

References

[1] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm . i. expected complexity," *IEEE transactions on signal processing*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[2] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE transactions on information theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[3] O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice code decoder for space-time codes," *IEEE communications letters*, vol. 4, no. 5, pp. 161–163, May 2000.

[4] L. M. Davis, "Scaled and decoupled cholesky and qr decompositions with application to spherical mimo detection," *Proc. IEEE WCNC*, pp. 326–331, Mar. 2003.

[5] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel," *Proc. ISSSE*, pp. 295–300, Sept. 1998.

[6] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in lattice, including a complexity analysis," in *Mathematics of Computation*, Apr. 1985, vol. 44, no. 170, pp. 463–471.

[7] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE transactions on signal processing*, vol. 53, no. 4, pp. 1474–1484, Apr. 2005.

[8] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm. ii. generalizations, second-order statistics, and applications to communications," *IEEE transactions on signal processing*, vol. 53, no. 8, pp. 2819–2834, Aug. 2005.

[9] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge Univ. Press, 2003.

[10] M. O. Damen, K.Abed-Meraim, and J.-C. Belfiore, "Generalized sphere decoder for asymmetrical space-time communication architecture," *Electronics letters*, vol. 36, no. 2, pp. 166–167, Jan. 2000.

[11] P. Dayal and M. K. Varanasi, "A fast generalized sphere decoder for optimum decoding of under-determined mimo systems," *in Proc. of 41st Annual Allerton Conf. on Comm. Control, and Comput.*, Oct. 2003.

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

[13] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 2nd ed. John Hopkins Univ. Press, 1989.

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, Sept. 2004.

[15] D. Pham, K. R. Pattipati, P. K. Willett, and J. Luo, "An improved complex sphere decoder for v-blast system," *IEEE signal processing letters*, vol. 11, no. 9, pp. 748–751, Sept. 2004.