REDUCED COMPLEXITY SPHERE DECODING USING A GEOMETRICAL APPROACH

M. Abbasi^{*}, Ali A. Tadaion^{*}, S. Gazor[†]

 * SSPR Lab, ECE Dept Yazd University, Yazd, Iran
 [†] ECE Dept, Queen's University, Ontario, Canada

ABSTRACT

In this paper we propose an algorithm with reduced complexity for the sphere detection (SD) which is used in multiple input multiple output (MIMO) detection algorithms without any performance degradation. The trade-off between the complexity and the bit error rate is a main challenge in wireless MIMO systems. The maximum likelihood (ML) detector considered as the optimum detector in the literatures. Since the complexity of the naive ML detectors is significantly high, the SD algorithms are proposed to lower the complexity. In this paper, we use the result of the geometrical decoder (GD) proposed in [8] which performs as the ML detector and has lower complexity than SD algorithm. We propose a method to further reduce the complexity of this SD algorithm. We show that the complexity is further reduced by almost 60%, i.e., the number of nodes visited by the proposed SD method is in average 60% less than that of the original one.

1. INTRODUCTION

The naive maximum likelihood (ML) detector is implemented by the exhaustive search over the whole symbol vector space and has the exponential order of complexity in terms of the number of and constellation points transmit antennas. The ML criterion is to find the solution which minimizes the Euclidean norm of the error. Although considered as optimum, this detector is not feasible in practice [1]. To reduce the complexity, many algorithms are proposed which have near to optimum performance.

One of the attractive MIMO detection algorithm is the SD algorithms which was first proposed by some mathematicians to solve the closest vector point problem in a lattice; this problem is equivalent to the ML detection of symbols in MIMO systems [2], [3]. The improved algorithm proposed by Finck and Pohst (FP-SD) efficiently enumerates all the lattice points within the hyper-sphere with certain radius [4]. The algorithm proposed by Schonorr and Euchner [5], (SE-SD) is more efficient since the search is performed in a zig-zag order within the sphere and the sphere radius is updated as a closer lattice point is found within the vicinity of to the input vector.

An alternative interesting geometrical approach in MIMO

detection is proposed first in [6] which performs as linear detectors using a post-processing algorithm. This method is extended to be employed for the higher dimensions and larger number of constellations [7]. Based on the geometrical approach in [8], an interesting detector is proposed with a reduced complexity and no compromise in performance.

In this paper, we propose a hybrid technique to efficiently reduce the complexity of the SD algorithm. We use geometrical interpretation of the ML criterion as [8] in order to limit the range of the search for each component of the unknown transmit vector. We apply this limitation in the FP-SD algorithm and iteratively form a new smaller search space within the search sphere of the FP-SD algorithm. We name this algorithm as geometrical sphere detection (GSD) and observe that it reduces the number of search vectors without making any compromise on the performance. Our simulation examples confirm the optimal performance of the proposed method and show a significant reduction in the complexity.

The remaining of the paper is organized as follows. In Section 2, we introduce the system model and the detection criterion. In Section 3, we summarize the FP-SD algorithm and review the geometrical detection algorithms in Section 4. Section 5 presents our proposed geometrical SD (GSD) algorithm. In Section 6, we provide some numerical results and compare the complexity of the proposed GSD algorithm with the conventional SD algorithms and confirm the optimality of the proposed algorithm. Finally in Section 7, we give some concluding remarks.

2. SYSTEM MODEL

We consider an uncoded MIMO system with m transmit and n receive antennas and we assume that $m \ge n$. The baseband received vector is modeled as [1]

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

where $\tilde{\mathbf{n}}$ is the complex additive noise vector of independent and identically distributed white Gaussian elements with zero-mean and variance σ^2 i.e. $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I), \, \tilde{\mathbf{y}} \in \mathbb{C}^{n \times 1}$ represents the received signal vector, \mathbb{C} is the set of complex numbers and $\tilde{\mathbf{x}} \in \mathbb{C}^{m \times 1}$ is the transmitted vector whose elements are chosen from a Pulse Amplitude Modulation (PAM) or Quadratic Amplitude Modulation (QAM) signal constellation. In this paper, we assume that the transmitter modulation is QAM. We assume that the channel state information (CSI) matrix $\tilde{H} \in \mathbb{C}^{n \times m}$ is perfectly known at the receiver and is quasi-static during the symbol frame. The CSI matrix \tilde{H} is assumed to be Rayleigh block flat fading channel, i.e., the set of the real and imaginary parts of all its elements are independent, Gaussian and have zero-mean and unit variance. For simplicity like [10], we transform the complex representation of (1) into the real representation:

$$\mathbf{y} = H\mathbf{x} + \mathbf{n}.\tag{2}$$

The dimensions are doubled in (2) compared with (1). The optimum detector for this model is the ML detector [11] and minimizes the squared norm of the error as follows

$$\hat{\mathbf{x}}_{\mathrm{ML}} = \underset{\mathbf{x} \in \chi^{2m}}{\mathrm{arg\,min}} \|\mathbf{y} - H\mathbf{x}\|^2, \tag{3}$$

where χ is the set of constellation points. Hereafter we refer to $\|\mathbf{y} - H\mathbf{x}\|^2$ as the squared error distance (SED). The naive ML algorithm is to search and evaluate the SED for all possible transmitted vectors and choose the one with the smallest SED. However the number of possible transmitted vectors is $|\chi|^{2m}$, where |.| denotes the set cardinality. The evaluation of the for one candidate requires an order of mn complex multiplications, i.e., the computational complexity of the naive ML is of order of $mn|\chi|^{2m}$ which makes it impractical as m and $|\chi|$ increase. In the following, we review FP-SD algorithm that efficiently reduces the search space.

3. SPHERE DETECTION ALGORITHMS

From a lattice view point, H is referred to as the generator matrix of the lattice defined by $\mathcal{L} = \{H\mathbf{x} | \mathbf{x} \in \chi^{2m}\}$ [12]. Therefore, the solution of (3) is the closest point of the lattice \mathcal{L} to the received vector \mathbf{y} . This problem is known as the closest lattice point problem and for arbitrary H has been shown to be NP-hard [4]. In sphere detection, instead of evaluating all the lattice points, only a subset of points are evaluated that are inside an sphere with a certain radius around the receive vector \mathbf{y} . According to [9], a latice point $H\mathbf{x}$ is said to be in the sphere of radius d_0 centered at \mathbf{y} where

$$\|\mathbf{y} - H\mathbf{x}\|^2 \le d_0^2. \tag{4}$$

Using QR-factorization of H, we can write H as

$$H = \begin{bmatrix} Q & P \end{bmatrix} \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix},$$

where R is an $2m \times 2m$ upper triangular matrix, **0** is an $2(n-m) \times 2m$ zero matrix and [Q P] is an $2n \times 2n$ unitary matrix. Using this factorization, it is easy write (4) [11]

$$\|\mathbf{y} - H\mathbf{x}\|^{2} = \|[Q P]^{T}\mathbf{y} - \begin{bmatrix} R \\ 0 \end{bmatrix} \mathbf{x}\|^{2}$$
$$= \|Q^{T}\mathbf{y} - R\mathbf{x}\|^{2} + \|P^{T}\mathbf{y}\|^{2}.$$

Thus we can rewrite (4) as follows

$$\|\mathbf{y}' - R\mathbf{x}\|^2 \le d_0^2 - |P^T\mathbf{y}|^2 \triangleq d^2,$$
(5)

where $\mathbf{y}' \triangleq Q^T \mathbf{y}$. We can rewrite (5) as

$$\sum_{i=1}^{2m} \left(y'_i - \sum_{j=i}^{2m} r_{ij} x_j \right)^2 \le d^2.$$
 (6)

The FP-SD algorithm starts from i = 2m to i = 1 and finds all possible symbols for each x_i using (6). Since R is an upper triangular matrix, we obtain a set of condition from 2m to 1. Each condition gives and specific interval for x_i for a given set of values of $x_{i+1}, ..., x_{2m}$. In i^{th} step, x_i must be in the interval [LB (x_i) , UB (x_i)] defined by

$$LB(x_i) = \lceil \frac{1}{r_{ii}} (y'_i - \sum_{j=i+1}^{2m} r_{ij} x_j - d_i^2) \rceil,$$
(7)

$$UB(x_i) = \lfloor \frac{1}{r_{ii}} (y'_i - \sum_{j=i+1}^{2m} r_{ij} x_j + d_i^2) \rfloor, \qquad (8)$$

and $d_i^2 = d^2 - \sum_{k=i+1}^{2m} (y'_k - \sum_{j=k}^{2m} r_{k,j}x_j)^2$ where [.] and [.] denote the smallest and the largest constellation point greater than or equal to its argument, respectively. The FP-SD algorithm makes a search set, examines each member in (3) and chooses the one with smallest SED. The FP-SD algorithm has much lower complexity than the naive ML. The complexity (as [9]) usually is measured in the term of the number of nodes that are visited.

4. GEOMETRICAL MIMO DETECTION

In this section, we review the geometrical approach suggested in [8]. The optimal ML solution in (3) is equivalent to [13]:

$$\hat{\mathbf{x}}_{\mathrm{ML}} = \operatorname*{arg\,min}_{\mathbf{x} \in \chi^{2m}} (\mathbf{x} - \mathbf{x}_c)^T H^T H(\mathbf{x} - \mathbf{x}_c), \qquad (9)$$

where $\mathbf{x}_c = (H^T H)^{-1} H^T \mathbf{y}$. By substituting $M = (H^T H)^{-1}$, we can rewrite (9) as:

$$\hat{\mathbf{x}}_{\mathrm{ML}} = \operatorname*{arg\,min}_{\mathbf{x} \in \chi^{2m}} f(\mathbf{x}),\tag{10}$$

where : $f(\mathbf{x}) \stackrel{\Delta}{=} (\mathbf{x} - \mathbf{x}_c)^T M^{-1} (\mathbf{x} - \mathbf{x}_c)$. Let us denote the eigenvalue decomposition of M as [6]:

$$M = V\Lambda V^T, \tag{11}$$

where $V = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{2m}], \{\mathbf{v}_i\}$ are the right singular vectors of $H, \Lambda = \text{diag}[\lambda_1, \cdots, \lambda_{2m}]$ is the diagonal matrix and $\{\lambda_i\}$ are the eigenvalues of M in ascending order (λ_i can be calculated from the singular values of H). In geometrical approach, the cross sections (contours) of the elliptic function $f(\mathbf{x})$ in (10) are hyper ellipsoid, i.e.,

$$f(\mathbf{x}) = a^2,\tag{12}$$



Fig. 1. Function $f(\mathbf{x})$ with three horizontal cross section and the projection of them in χ^{2m} space.

is a hyper ellipsoid as shown in Figure 1 centered at \mathbf{x}_c , where a > 0. The axis of these ellipsoids are $\{\mathbf{v}_i\}$ with diameters $\{a\sqrt{\lambda_i}\}$.

Figure 1 shows this elliptic paraboloid and three cross sections at three heights a_1^2, a_2^2 and a_3^2 . All these hyper-ellipsoids are centered at \mathbf{x}_c where $f(\mathbf{x})$ takes its minimum value at \mathbf{x}_c $(f(\mathbf{x}_c) = 0)$. Since (9) is equivalent to (3), the optimal solution of (9) in χ^{2m} also minimizes (3). Moreover, the function $f(\mathbf{x})$ is a convex in \mathbf{x} . Thus any hyper-ellipsoid containing one point of χ^{2m} on its boundary must contain the optimal ML solution. Therefore, we only need to evaluate those points of χ^{2m} which are within such an hyper-ellipsoid, i.e., all we have to do is to find a hyper-ellipsoid (as small as we can) which contains at least one point from χ^{2m} . For example in Figure 1, the three hyper-ellipsoids contain points from χ^{2m} , however, hyper-ellipsoid 1 is preferred as it is the smallest one. Our optimization problem is to find the smallest hyperellipsoid which is NP-hard as the smallest one contains only the optimal solution. However, we can start and choose a larger hyper-ellipsoid by using any point from χ^{2m} . To make the hyper-ellipsoid as small as we can use the output of an inexpensive detector such as the linear detector. Then, we aim to search only over a subset points from χ^{2m} which belongs to this hyper-ellipsoid and choose the point which minimizes (9). In [8], it is proposed to search over a hyper-rectangle enclosing this hyper-ellipsoid. They propose to use \mathbf{x}_{ZF} in order to calculate $a_{\rm ZF} = \sqrt{f(\mathbf{x}_{\rm ZF})}$. They argue that the *i*th component of x must be the following interval:

$$x_{i}^{\min} = x_{c_{i}} - \sum_{j=1}^{2m} |v_{j,i}a_{\rm ZF}\sqrt{\lambda_{i}}|, \qquad (13)$$

$$x_i^{\max} = x_{c_i} + \sum_{j=1}^{2m} |v_{j,i} a_{\text{ZF}} \sqrt{\lambda_i}|.$$
 (14)

Thereby, they proposed to search over the above hyper-

rectangle. In the next section, we propose to employ the intersection of this hyper-rectangle and χ^{2m} as an input to the SD algorithm in order to substantially reduce the computational complexity of the SD algorithm.

5. THE PROPOSED ALGORITHM

From (13) and (14), we know that we must have $x_i \in [x_i^{\min}, x_i^{\max}]$ [8]. We here propose to combine the SD algorithm in the Section 3 with this hyper-rectangle constraint and refer to this proposed MIMO detection as Geometrical SD (GSD) algorithm. The hyper-rectangle in (13) and (14) defines new 2m constraints which reduces the search space without loss of optimality. In addition, the constraints in (7) and (8) can be also used to further reduce the search space according to the SD algorithm. In the GSD algorithm, we use the intersection of these conditions, i.e., we substitute the following in the SD algorithm:

$$\overline{\text{LB}}(x_i) = \max\{\lceil x_i^{\min} \rceil, \lceil \frac{1}{r_{ii}} (y_i' - \sum_{j=i+1}^{2m} r_{ij} x_j - d_i^2) \rceil\},$$
(15)

$$\overline{\mathrm{UB}}(x_i) = \min\{\lfloor x_i^{\max}\rfloor, \lfloor \frac{1}{r_{ii}}(y_i' - \sum_{j=i+1}^{2m} r_{ij}x_j + d_i^2)\rfloor\}.$$
(16)

The search space is reduced into the intersection of:

- 1. a hyper-sphere as given in the SD algorithm,
- 2. and a hyper-rectangle which is an enclosure of a hyperellipsoid with a radius that can be reduced as a better solution is found during the search process.

The computational cost is substantially reduced as our proposed method searches over a substantially smaller subset of χ^{2m} . The GSD algorithm is more efficient than both since the intersection of two is always smaller than the each one.

The optimality of the propose algorithm is not compromised and in all our simulations as it is guaranteed to find the solution of the SD algorithm, i.e., the bit error rate (BER) of the GSD is identical to that of the proposed algorithm. The constraints (13) and (14) lead to the ML solution [8], moreover (7) and (8) lead to the same solution [11], therefore the intersection of them also allows to find the ML solution. The proposed method could be employed in combination with the FP-SD,therefore We summarise the GSD algorithm in the following steps:

- 1) Initialize $\mathbf{x} \leftarrow \mathbf{x}_{ZF}$ using ZF detector.
- 2) Calculate the radius of the sphere $d^2 \leftarrow \|\mathbf{y} H\mathbf{x}\|^2$
- 3) Calculate the radius hyper-ellipsoid: and $a^2 \leftarrow f(\mathbf{x})$.
- 4) Calculate $\overline{\text{LB}}_i$ and $\overline{\text{UB}}_i$ using 15 and 16 for i = 1, ..., 2m. 5) Employ the FP-SD algorithm using $x_i \in [\overline{\text{LB}}_i, \overline{\text{UB}}_i]$ as the search interval.

6) Evaluate (3) for members of χ^{2m} satisfying constraint in step 4 and 5, and keep x if (3) is smaller than the previously

smallest value.

6) (optional) Go to Step 2 if \mathbf{x} is not the same as the previous solution. In the table 1, the FP-GSD algorithm is summarized in more details.

Table 1. The proposed method for FP-SD (FP-GSD algorithm)

Input:	$Q, P, R, \mathbf{y}, \mathbf{y}', d_0$ (d_0 is the SED of ZF solution)
Step1)	for $i = 1,, 2m$ set x_i^{\min} and x_i^{\max} according to
	equation (13) and (14).
Step 2)	Set $i = m$, $d_m^2 = d_0^2 - P^T \mathbf{y} ^2$, $\mathbf{y}'_{m m+1} = \mathbf{y}'_m$,
Step 3)	Set UB(x_i) = min{ $\lfloor x_i^{\max} \rfloor$, $\lfloor (d'_i + y_{i i+1})/r'_{i,i} \rfloor$ },
	$x_{i} = \max\{\lceil x_{i}^{\min} \rceil, \lceil (-d_{i}' + y_{i i+1})/r_{i,i} \rceil\} - 1$
Step 4)	Set $x_i \leftarrow x_i + 1$. If $x_i < UB(x_i)$, go to step 6;
	else go to step 5.
Step 5)	i = i + 1; if $i = 2m + 1$ end of the algorithm,
	choose \mathbf{x} whit the smallest SED as the solution;
	else go to step 4.
Step 6)	If $i = 1$ go to step 7 else $i = i - 1$, $y'_{i i+1} =$
	$y'_{i} - \sum_{j=i}^{2m} r_{ij} x_{j}$ and $d'^{2}_{i+1} = d'^{2}_{i} - y_{i+1 i+2} - d'^{2}_{i+1 i+2}$
	$r_{i+1,i+1}x_{i+1}$ and go to step 2.
Step 7)	A new vector is found. Save the x and its SED,
	$d_{2m}^{\prime 2} = d_1^{\prime 2} + (y_1^{\prime} - r_{1,1}x_1)^2$ and go to step 4.

6. SIMULATION RESULTS

In this section, we use simulations to illustrate the computational complexity reduction in terms of average number of nodes visited by GSD algorithm. The optimality is measured by bit error rate (BER) using monte carlo simulations and the complexity (like [9]) is shown by the average number of nodes visited in the algorithm. We have repeated our simulations for 4×10^4 independent channel realizations and then averaged the resulting BER and the number of visited nodes. Figure 2 shows that the BER of the proposed GSD algorithms, the traditional SD algorithm and the ZF method for two 4×4 and 8×8 MIMO systems. This figure shows that these algorithms have identical BER performances as they are all optimal except the ZF method which shows a considerable performance gap with the optimum detectors. For original SD algorithm, the initial search radius is set to the SED of the ZF solution.

Table 2 show the average number of nodes visited by the original FP-SD and FP-GSD (we briefly write it GSD) algorithms with different settings. These numbers represent the relative computational complexity of these detectors. As we see this Table represent the complexity of a 4×4 MIMO system with 16 and 64 QAM modulations at five different SNRs. From this table, we see that the GSD algorithm has a significantly reduced computational cost. For example for the 16-QAM constellation, this reduction is 56% and 70% at SNR = 25 and SNR = 30dB, respectively. For a 64-QAM constellation, the reductions are 49% and 62% respectively at



Fig. 2. BER comparison of different SD algorithm.

 Table 2.
 Complexity comparison for MIMO system in terms of the average number of nodes visited

4×4 system						
SNR	15	20	25	30		
16 QAM SD	266	149	79	54		
16 QAM GSD	224	96	34	16		
complexity reduction	15%	35%	56%	70%		
64 QAM SD	4056	1925	727	477		
64 QAM GSD	3632	1487	369	177		
complexity reduction	10%	22%	49%	62%		
8×8 system						
SNR	25	30	35	40		
16 QAM SD	3512	1437	650	197		
16 QAM GSD	2838	867	294	66		
complexity reduction	19%	39%	54%	66%		

the same SNRs. Table 2 also confirm similar results for 8×8 MIMO system with 16-QAM modulation.

7. CONCLUSION

In this paper employing a geometrical approach, we proposed a algorithms to reduce the complexity of the SD algorithm for MIMO detection. Using mont carlo simulation, we showed that proposed method results in significant reduction in the average computational cost. This reduction for 4×4 and 8×8 MIMO system was almost 70% and 60% respectively.

8. REFERENCES

 E. Biglieri, R. Calderbank, A. constantinides, A. goldsmith, A. paulraj and H. V. Poor, *MIMO Wrieless Communication*. New York, NY, USA: Cambridge University Press, 2007.

- [2] A. Burg, M. Borgmann, M. Wenk, M. Zellweger, W. Fichtner, and H. Bolcskei, "VLSI implementation of MIMO detection using the sphere decoding algorithm," *Solid-State Circuits, IEEE Journal of*, vol. 40, pp. 1566-1577, 2005.
- [3] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *Information Theory, IEEE Transactions on*, vol. 45, pp. 1639-1642, 1999.
- [4] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Computation*, vol. 44, pp. 463-471, Apr. 1985.
- [5] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," Math. Programming, vol. 66, pp. 181191, 1994.
- [6] H. Artes, D. Seethaler, and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection," *Signal Processing, IEEE Transactions on*, vol. 51, pp. 2808-2820, 2003.
- [7] M. Samuel and M. P. Fitz, "Geometric Decoding Of PAM and QAM Lattices," in Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE, 2007, pp. 4247-4252.
- [8] Z. Y. Shao, S. W. Cheung, and T. I. Yuk, "A Simple and Optimum Geometric Decoding Algorithm for MIMO Systems," in Wireless Pervasive Computing, 2009. ISWPC 2009. 4th International Symposium on, 2009, pp. 1-5.
- [9] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *Signal Processing*, *IEEE Transactions on*, vol. 53, pp. 2806-2818, 2005.
- [10] H. Shuangshuang and C. Tellambura, "A Complexity-Efficient Sphere Decoder for MIMO Systems," in Communications (ICC), 2011 IEEE International Conference on, 2011, pp. 1-5.
- [11] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *Information Theory, IEEE Transactions on*, vol. 49, pp. 2389-2402, 2003.
- [12] D. Wübben, D. Seethaler, J. Jaldén, and G. Matz, "Lattice reduction: A survey with application in wireless communications," *Signal Processing Magazine, IEEE*, vol. 28, pp. 70-91, 2011.

- [13] L. Hanzo,Y. Akhtman, L. Wang,M. Jiang, MIMO-OFDM for LTE, Wi-Fi and WiMAX, Coherent versus Non-coherent and Cooperative Turbo-transceivers Southampton, UK: Wiley Press, 2011.
- [14] M. Abbasi, A. A. Tadaion, and M. R. Taban, "An improvement of ZF/LR method in MIMO detection," *in Telecommunications (IST), 2012 Sixth International Symposium on*, 2012, pp. 426-430.