# A POLYNOMIAL TIME ALGORITHM FOR EXACT MAXIMUM LIKELIHOOD DECODING OF MIMO CHANNELS: A DISCRETE GEOMETRIC APPROACH

*Ejaz Khan, Dirk Slock*

Mobile Communications Department
Institut Eurécom, B.P. 193, 06904 Sophia–Antipolis CEDEX, France
email: Ejaz.Khan@eurecom.fr

## ABSTRACT

The information capacity of wireless communication systems may be increased dramatically by employing multiple transmit and receive antennas. Often the optimal receiver in Maximum likelihood sequence detector (MLSD), which is considered to be practically infeasable due to high computational complexity. Therefore, in practice one often settles with less complex suboptimal receiver structure. In this chapter, we propose a polynomial time algorithm for exact maximum likelihood (ML) decoding for MIMO channels. The problem is posed as maximizing a quadratic form in $N$ binary variables (BPSK case) with the vertices of a hypercube as constraint. We consider $M$ receive antennas and $N$ transmit antennas. We assume that $M < N$, and that $M$ is fixed. The maximization of ML cost function with vertices of hypercube, i.e., $\{-1, 1\}^N$, as constraints, translates to having a symmetric matrix in quadratic form with fixed rank, $M$ and with the hypercube constraint. With singular value decomposition (SVD) of the symmetric matrix and suitable affine transformation of the hypercube constraint one ends up with maximizing a quadratic function (Euclidean distance) over extreme points (vertices) of zonotope (definition of zonotope will be given later). Using a classical theorem of discrete geometry, it is shown that the vertices search can be done in polynomial time, $O(N^{M-1})$. The overall complexity of the algorithm is the complexity to find extreme point of zonotope plus the complexity of the SVD operation plus the evaluation of the objective function at the vertices. To find the vertices of zonotopes, an efficient algorithm called reverse search algorithm can be employed [1,6,7]. Our approach has potential benefits over currently popular sphere decoding scheme [3]. The average case complexity of sphere decoding scheme is $O(N^3)$ plus the complexity to perform QR decomposition $(\frac{2}{3}N^3)$ of the channel, when radius, r, is correctly chosen (which is NP-hard problem). Also at low SNRs the complexity of the sphere decoder explodes. The other problem with sphere decoding is that some form of the heuristics is used to choose the radius of hypersphere. On the contrary, the proposed method has polynomial complexity independent of SNR and also no heuristic is used in the algorithm.

## 1. INTRODUCTION

Multiple antenna wireless communication systems are capable of providing data rate at potentially very high rates. To secure high reliability of the data transmission special attention has to be payed to the design of the receiver. In many communications systems the optimal receiver structure is maximum likelihood sequence detector (MLSD). However, computational complexity of the traditional MLSD often prohibits its practical implementation. Thus often one settles with suboptimal receivers like, MIMO-DFE, BLAST, are some of them. Recently, sphere decoding has gained quite popularity due to its average polynomial complexity (at high SNR) in the number of variables (antennas). The sphere decoder has average complexity of $O(N^3)$, when the radius of the hypersphere is optimally chosen (which is NP-hard). But it has exponential complexity for low SNRs. The other problem with the sphere decoder is that at each time instant it has average complexity, $O(N^3)$, which can be computationally very complex for large tramsmitted data blocks. In this paper, we focus on MLSD. It is assumed that the receiver has perfect knowledge of the channel. We propose a novel exact method for data detection using some beautiful results in discrete geometry. The detection is based on maximizing ML function subject to vertices of zonotope (special polytope) constraint. We will show that optimal solution to the problem is polynomially bounded in $N$.

## 2. SIGNAL MODEL

We assume a discrete time block fading multiple antenna channel model with $N$ transmit and $M$ receive antennas, we have assumed that the receiver has perfect channel knowl-

Eurécom's research is partially supported by its industrial partners: Ascom, Cégétel, France Télécom, Hitachi, IBM France, Motorola, Swisscom, Texas Instruments, and Thomson CSF

edge. The received signal at an instant is

$$x = Hs + n \quad (1)$$

where

$$H \in R^{M \times N}, \quad (2)$$

is a known channel matrix,

$$n \in R^{M \times 1} \quad (3)$$

is the i.i.d. zero mean white Gaussian noise with variance $\sigma^2$. The fading coefficients are i.i.d. Gaussian with zero mean and unit variance. Under the aforementioned assumptions the ML criterion requires us to find $s \in R^{N \times 1}$, where $s_i \in \{-1, 1\}$, which minimizes $||x - Hs||^2$. The problem can be written (after neglecting constant terms) as

$$f(s) = max_{s \in \{-1,1\}^N} \;\; s^T Js + 2c^T s, \quad (4)$$

where

$$J = -H^T H \;\; and \;\; c = Hs. \quad (5)$$

The above equation, i.e., $f(s)$, is equivalent to

$$f(s) = max_{y \in \{-1,1\}^N, \; y_{N+1}=1} \;\; y^T \bar{L} y, \quad (6)$$

where $\bar{L}$ is

$$\bar{L} = \begin{pmatrix} J & c \\ c^T & 0 \end{pmatrix} \quad (7)$$

Since this cost function is symmetric, $y_{N+1} = 1$ need not been maintained explicitely.

## 3. RELATIONSHIP BETWEEN THE QUADRATIC PROGRAM AND MLSD

In the following lines, we show the relationship between quadratic $(0-1)$ programming (QP) problem and MLSD. The $(0-1)$ quadratic problem can be written as

$$QP = min_{x \in \{0,1\}^n} \;\; x^T Qx \quad (8)$$

Where Q is real symmetric positive matrix. With the change of variable,

$$y = \frac{s+e}{2}, \quad (9)$$

in eq (4),where

$$s \in \{-1, 1\}^N \;\; and \;\; e = [1, 1 \ldots, 1]^T \quad (10)$$

$e$ is all ones vector, the objective function can be written as

$$\Psi(y) = (2y - e)^T J(2y - e) + 2c^T (2y - e). \quad (11)$$

It is straight forward to see that the above objective function can be written as

$$\Psi(y) = max_{y \in \{0,1\}^N} \;\; y^T \bar{J} y + 2y^T \bar{c}, \quad (12)$$

where

$$\bar{J} = 4J \;\; and \;\; \bar{c} = -2Je + 2c. \quad (13)$$

$\Psi(y)$ can be further written as

$$\Psi(y) = max_{y \in \{0,1\}^N, \; y_{N+1}=1} \;\; y^T \tilde{J} y \quad (14)$$

where

$$\tilde{J} = \begin{pmatrix} \bar{J} & \bar{c} \\ \bar{c}^T & 0 \end{pmatrix} \quad (15)$$

Since the cost function in eq (14) is symmetric. $y_{N+1}$ need not be maintained explicitely. Therefore the inclusion of the linear term leads to the problem in eq (14) of size increased by one. Hence we have shown that our detection problem is equivalent to $(0-1)$ quadratic problem. We will use eq (6) for analysis because the constraint $\{-1, 1\}^N$(i.e, hypercube), has symmetry property which can later be exploited.

## 4. DISCRETE GEOMETRIC APPROACH TO MLSD

A basic problem in discrete optimization consists in optimizing a quadratic over some hypercube. This type of problem is NP-hard, and it is still considered a computational challenge to solve general modest size problems of this type to optimality. Quadratic programming (QP) over vertices of cube appears in various equivalent formulations in the literature. Our problem (MLSD) is maximization of a quadratic function over vertices of hypercube. Before delving into the solution of this problem we define some geometrical objects.

Polytope:

A polytope (convex polytope) is a convex hull of finite set of points in $R^d$ (which are always bounded) or as bounded intersection of finite set of half spaces. Polytope can also be defined as a finite region of d-dimensional space enclosed by a finite number of hyperplanes.

Zonotope [2, pp. 198-199]:

Zonotopes are special polytopes that can be viewed in various ways : for example, as projections of cubes, as Minkowski sums of line segments, and set of bounded linear combinations of vector configurations. Each of these description gives different insight into the combinatorics of zonotopes. A zonotope is the image of a cube under an affine projection, $Z \subseteq R^d$ of the form [2],

$$Z = Z(V) = VC_p + z = \{Vy + z : y \in C_p\} \quad (16)$$

$$Z = Z(V) = \{x \in R^d : x = z + \sum_{i=1}^{p} x_i v_i, \quad -1 \leq x_i \leq 1\} \tag{17}$$

for some matrix (vector configuration),

$$V = [v_1, \dots, v_p] \in R^{d \times p}. \tag{18}$$

Equivalently, since every d-cube $C_d$, is a product of line segments

$$C_d = C_1 \times \dots \times C_1, \tag{19}$$

we get that every zonotope is the Minkowski sum of a set of line segments. Infact, if $\pi$ is linear we get

$$\begin{aligned} Z(V) &= \pi(C_1 \times \dots \times C_1) \\ &= \pi(C_1) + \cdots + \pi(C_1) \\ &= [-v_1, v_1] + \dots [-v_p, v_p] \end{aligned} \tag{20}$$

and thus

$$Z(V) = [-v_1, v_1] + \cdots + [-v_p, v_p] + z, \tag{21}$$

for an affine map given by

$$\pi(y) = Vy + z. \tag{22}$$

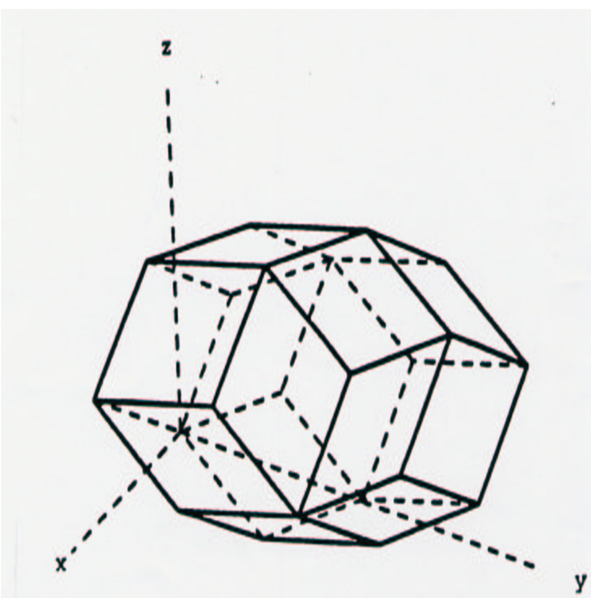


**Fig. 1**. 3 dimentional zonotope with 5 generators.

Having defined polytope/zonotope, we can proceed with our problem. First, we begin by spectral factorization of

$$\bar{L} = UU^T, \tag{23}$$

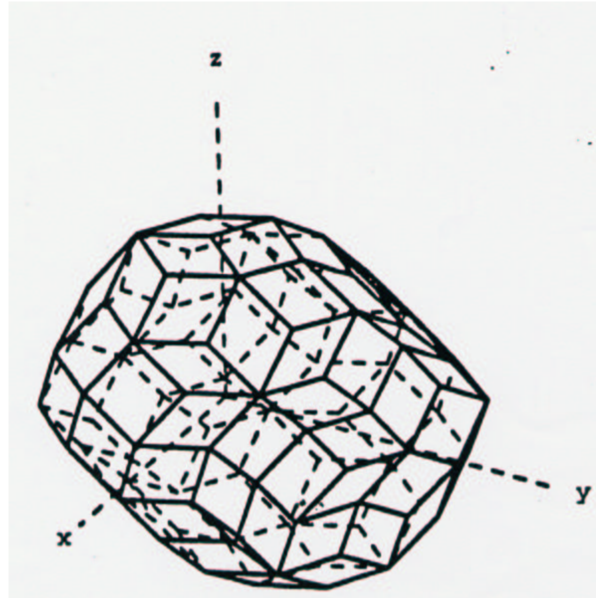


**Fig. 2**. 3 dimensional zonotope with 10 generators.

where $U \in R^{M \times N}$ is the matrix composed of suitably scaled eigenvectors. We can write eq (6) as

$$\begin{aligned} \Psi(y) = y^T \bar{L} y = ||U^T y||^2 \\ subject\ \ to\ \ y \in \{-1, 1\}^{N+1} \end{aligned} \tag{24}$$

Consider affine map [2 pp. 199],

$$R^N \longrightarrow R^M : \ \ z = U^T y. \tag{25}$$

This linear transformation maps a $\{-1, 1\}^{N+1}$ (hypercube) into a symmetric zonotope. For every extreme point $\bar{z}$ of zonotope there exists an extreme point $\bar{y} \in \{-1, 1\}^{N+1}$ such that $\bar{z} = U^T \bar{y}$ and thus eq (6) can be written in the following form

$$\Psi(y) = max_{z \in Z_{extreme}} ||z||^2 \tag{26}$$

From the above equation it is clear that our objective function and constraint both are symmetric and some of the extreme points of hypercube will correspond to some points lying inside or on the facets of zonotope. Observe that extreme points which lie inside or on the facets of the zonotope cannot be candidate for the maximization of our objective function. Therefore the maximum is attained at some vertex $\bar{z}$ of $Z$. Thus MLSD is thus reduced to the enumeration of vertices of zonotope $Z$. Now the problem is to calculate the number of facets and vertices of a $M$ dimensional zonotope given by $N+1$ generators. Let $f_0(Z)$ and $f_{M-1}(Z)$ denote the number of vertices (extreme points) and facets of $Z$, repectively. The answer to the above question is given by the following classical theorem in discrete geometry:

**Theorem**:
Let $Z$ be $M$ dimensional zonotope given by $N+1$ generators $(N > M)$. Then

$$f_0(Z) \leq 2 \sum_{i=0}^{M-1} \left( \begin{array}{c} N \\ i \end{array} \right) \quad , f_{M-1}(Z) \leq 2 \left( \begin{array}{c} N+1 \\ M-1 \end{array} \right) \quad (27)$$

where

$$\left( \begin{array}{c} m \\ n \end{array} \right) = \frac{m!}{(m-n)!n!} \quad (28)$$

Furthermore, the equalities are attained by certain zonotopes and therefore the bounds are best possible. From the above theorem it is clear that the upper bound of $f_{M-1}$ is $O((N+1)^{M-1}) \approx O(N^{M-1})$, for large $N$. The bound on the number of vertices is the expression

$$f_0(Z) \leq 2 \sum_{i=0}^{M-1} \left( \begin{array}{c} N \\ i \end{array} \right) \quad (29)$$

and the dominating one is the last term

$$\left( \begin{array}{c} N \\ M-1 \end{array} \right) \quad (30)$$

which is of $O(N^{M-1})$ for large $N$. Thus the number of vertices are polynomially bounded and there exists efficient algorithm to generate extreme points of $Z$ as given by the following theorem.

**Theorem**[5] :
Given $N+1$ generators of a zonotope there is $O((N+1)^M) \approx O(N^M)$, for large $N$, time algorithm to generate extreme points of zonotope for $M \geq 2$.

Uptil now we know the bound on the vertices but we do not know an algorithm to generate them. In order to explain it, we need a relationship between arrangements of hyperplanes and zonotopes.

Arrangements of hyperplanes [4, pp.4]:
A finite set of hyperplane in $E^d$ defines a dissection of $E^d$ into connected pieces of various dimensions. We call this dissection the arrangement $A(H)$ of H. For example, a finite set of lines in two dimensions dissects the plane into connected pieces of dimensions two, one and zero. It has been shown [4, pp. 20-26] that a zonotope in $E^d$ corresponds to an arrangement in $E^{d-1}$. For example a two dimensional zonotope has corresponding one dimensional arrangement. In [4,5] an algorithm is given to construct arrangements. The overall structure of the algorithm is given below.

The overall structure:
The construction of an arrangement proceeds incrementally, that is, the arrangement is built by adding hyperplane one at a time to the already existing arrangement. The order in which the hyperplanes are added is irrelevant. Let $H$ denote the set of hyperplanes $H = [h_1, \ldots, h_n]$ in $E^d$ and define

$$H_i = \{h_1, \ldots, h_i\} \quad (31)$$

for $1 \leq i \leq n$. $D(H)$ denotes the data structure to be described that represents the arrangement $A(H)$. It is assumed that the normal-vector of hyperplanes in $H$ span $E^d$. Let the normal vectors of $h_1, \ldots, h_d$ span $E^d$. Construct $D(H_d)$. For i running from $d+1$ to $n$, construct $D(H_i)$ from $D(H_{i-1})$ by insertion of $h_i$. Finally, $D(H) = D(H_n)$. Unfortunately, this algorithm may not be very practical because it has to store in memory the list of all vertices and faces generated before. This means that only the storage of vertices is of size $O(N^{M-1})$. In order to alleviate the complexity there exits an efficient algorithm, known as reverse search algorithm [1,6,7], for generating full dimensional regions. The advantage of this algorithm is that it can be highly parallelized and is also space and time efficient. In order to explain the basic idea of reverse search, let $G$ be a connected graph and suppose we have some objective function to be maximized over these vertices. A local search algorithm on $G$ is a deterministic procedure to move from any vertex to some neighboring vertex which is larger with respect to objective function until there exists no better neighboring vertex. A vertex without a better neighboring vertex is called local optimal. The algorithm is finite for any starting vertex, it terminates in finite number of steps. Simplex algorithm is an example of local search algorithms.

Let us imagine the simple case that we have finite search algorithm and there is only one local optimum vertex $x^*$ ( which is also optimal solution). Consider the directed graph $T$ with same vertex set as $G$ and the edges which are all ordered pair $(x, x^{'})$ of consecutive vertices $x$ and $x^{'}$ generated by local search algorithm. It should be clear that $T$ is a tree spanning all vertices with the only sink $x^*$. Thus if we trace this graph $T$ from $x^*$ systematically, say by depth first search, we can enumerate all vertices. The major operation here is tracing each edge against its orientation which corresponds to reversing the local search, while the minor work of backtracking is simply performing the search algorithm itself. We do not have to store any information about visited vertices for this search because $T$ is itself a tree. Observe that for each vertex $x$, every vertex $y$ below $x$ in $T$ (those $y$ such that there is directed path from $y$ to $x$) has no larger objective value and whenever it detects a vertex with lower objective value, then abandon going lower in the tree. The advantage of the reverse search algorithm are:

1) Time complexity is proportional to the size of the output times a polynomial in the size of input
2) Space complexity is polynomial in the size of input,
3) Parallel implementation is straight forward. We believe that we can exploit the symmetry of our problem to further reduce the complexity, i.e., time and memory, of the reverse

search algorithm. Having found the vertices with the help of reverse search algorithm, the task remaining is to calculate the value of the objective function at those vertices. We need to evaluate the objective function only on half the number of vertices (thanks to symmetry of our constraint) resulting in overall complexity of $O(N^M) + 6N^3$ for the proposed method. $6N^3$ is the complexity to calculate SVD of a symmetric $N \times N$ matrix. Having found the vertex of zonotope that maximizes our objective function, the corresponding vertex of hypercube can be found by, $z = U^T y$.

## 5. CONCLUSIONS

In this paper, we have shown theoretically a polynomial time algorithm to decode exactly a MIMO system when the number of receiving antennas is fixed. By posing the problem as maximization of quadratic form over zonotope and using some classical theorems of discrete geometry, we were able to solve the problem in polynomial time. Comparing our method with the sphere decoding, we have the following three advantages over the later:

1) The sphere decoder has exponential complexity at low SNRs while there is no such problem in our method (it is independent of SNR).

2) The sphere decoder has polynomial complexity (at high SNRs) at each time instant (because the received signal point moves from one point to another in lattice at each time instant) for flat fading or for block fading channels, while no such problem exists in our method.

3) No heuristic is employed in our algorithm, where as in the sphere decoder radius of the sphere is chosen heuristically and choosing the optimum radius of the sphere is NP-hard.

The disadvantage of our algorithm is that the degree of the polynomial increases as the number of receive antennas, $M$, increases. Although, we have assumed perfect channel knowledge but the same analysis applies for noisy channel estimates too.

## 6. REFERENCES


[1] D. Avis, ”Tutorial on reverse search with C implementation ”, Tech. report, July 13, 2000.

[2] G.M. Ziegler, ”Lectures on polytopes”, Springer-Verlag, Graduate text in mathematics, 1995.

[3] H. Vikalo and B. Hassibi, ”Maximum likelihood sequence detection for multiple antenna system over dispersive channels via sphere decoding ”, Eurasip Journal on App. Signal Processing , no.5, pp. 525-531, May, 2002.

[4] H. Edelsbrunner, ”Algorithms in combinatorial geometry ”, Springer-Verlag, monographs on theoretical computer science, 1987.

[5] H. Edelbrunner, ”J. O’Rourke, and R, Seidel ”Constructing arrangements of lines and hyperplanes with applications ”, SIAM journal of Comput. no.15, pp. 341-363, 1986.

[6] D. Avis, and K. Fukuda, ”Reverse search for enumeration ”, Discrete app. mathematics pp. 21-46, 1996.

[7] D. Avis, ”A revised implementation of the reverse search vertex enumeration algorithm”, editors, in Polytopes-combinatorics and computation, pp. 177-198, Birkhauser, 2000.