A New Ordering for Efficient Sphere Decoding

Karen Su and Ian J. Wassell
Laboratory for Communication Engineering
Cambridge University Engineering Department
Trumpington Street, Cambridge, CB2 1PZ
Tel: +44 (0)1223 767 {021, 026}
Email: {ks349, ijw24}@eng.cam.ac.uk

***Abstract*— This paper presents a novel pre-processing stage that offers significant improvement in the computational efficiency of sphere decoding by imposing a geometrically-inspired *ordering* on the columns of the channel matrix. By studying the performance of a *genie* decoder, which has knowledge of the optimal radius, we find that the optimal ordering depends not only on the channel matrix, but also on the received point. Analysis of this idealized problem leads to the proposal of an enhanced ordering. We demonstrate via simulation that it closely matches with the optimal ordering and more importantly that it results in a dramatic increase in sphere decoding efficiency over a 4×4 MIMO flat fading channel. We emphasize that the performance benefit is particularly great at low SNRs and for high modulation orders, two traditionally challenging regimes for sphere decoders. We conclude by briefly discussing the polynomial complexity of the new ordering algorithm.**

***Index Terms*— Sphere decoding, nearest lattice point search, maximum-likelihood detection, (linear) MIMO systems, pre-processing, ordering.**

I. Introduction

The Sphere Decoder (SD) is a powerful approach to the Maximum-Likelihood (ML) detection of signals observed at the output of overdetermined Multiple Input Multiple Output (MIMO) systems whose inputs are drawn from a lattice alphabet [1]. It is based on the enumeration of lattice points located within a hypersphere of some radius centered at a *target*, e.g., the received signal point.

The Fincke-Pohst (F-P) and Schnorr-Euchner (S-E) strategies are two computationally efficient means of realizing this enumeration [2], and so they have come to form the foundation of many existing sphere decoders [3], [4]. Underlying both is the application of the QR factorization to the channel matrix. The upper triangular structure of the factored matrix enables the decoder to decompose the overall enumeration task into a set of smaller sub-tasks of decreasing dimension.

Even when using the more efficient S-E enumeration, it is known that the computational cost of sphere decoding is highly sensitive to the ordering of the columns of the channel matrix [2]. Unlike previous proposals, where ordering decisions are based only on the channel matrix [2], we formally show that the optimal ordering for sphere decoding depends on the channel matrix and also on the received point. This philosophy was also recently employed in [5], [6], however our work is distinguished from these by its geometric perspective and further improved decoder efficiency.

We begin our presentation in Section II with a precise formulation of the problem and some mathematical preliminaries. Next, in Section III we study the behaviour of a *genie* SD under different orderings of a two-dimensional example channel. This analysis leads to the design of the new ordering scheme, which is formalized in Section IV. Simulation results reported in Section V demonstrate its effectiveness in reducing the computational cost incurred by a standard SD. Finally we conclude with comments on the complexity of the ordering scheme in Section VI.

II. Preliminaries

In this paper we consider problems that can be modelled as the minimization of the squared Euclidean distance metric to a target point $\mathbf{r}$ over an M-dimensional finite discrete search set $\mathcal{X}^M \subset \mathbb{Z}^M$ with $|\mathcal{X}| = B$:

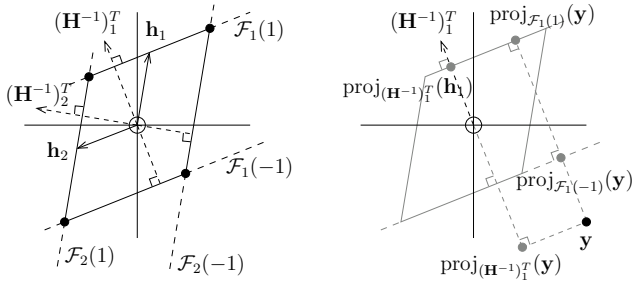
$$\mathbf{s}_* = \underset{\mathbf{s} \in \mathcal{X}^M}{\operatorname{argmin}} |\mathbf{r} - \mathbf{H}\mathbf{s}|^2, \tag{1}$$

where $\mathbf{r} \in \mathbb{R}^N$, $\mathbf{H} \in \mathbb{R}^{N \times M}$, and the optimization variables are the elements of $\mathbf{s}$.[1] Examples of such problems include ML detection of lattice coded signals and QAM modulated signals transmitted over MIMO flat fading channels, frequency selective fading channels, or multi-user channels.

We assume an overdetermined problem, i.e., that $M \leq N$, and that $\mathbf{H}$ is of full rank M. For communication over MIMO flat fading channels, this assumption means that there are at least as many receive (N) as transmit (M) antennas. We make use of the following notational conveniences: Given a matrix $\mathbf{A}$, let $\mathbf{a}_i$ denote the i^{th} column vector and $\mathbf{A}_{\setminus i}$ the tall submatrix comprised of all columns but the i^{th}. Given a vector $\mathbf{z}$, let z_i denote the i^{th} element and $\mathbf{z}_{\setminus i}$ the vector comprised of all elements but the i^{th}. We also denote by $\mathbf{e}_i$ the i^{th} elementary vector and by $\mathcal{I}$ the index set $\{1, 2, \ldots, M\}$. To distinguish between a variable itself and its value, we use the underline notation $\underline{s}_i$ or $\underline{\mathbf{s}}$ to refer to a variable, and $s_i \in \mathcal{X}$ or $\mathbf{s} \in \mathcal{X}^M$ to indicate a particular value.

This work was supported by Universities UK, the Cambridge Commonwealth Trust, the Natural Sciences and Engineering Research Council of Canada and Trinity Hall.

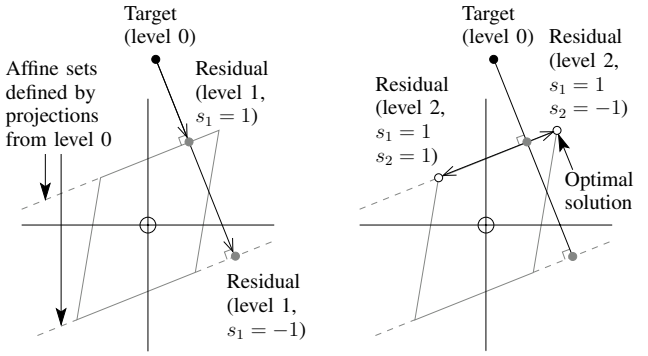
[1] The complex case where $\mathbf{s} \in (\mathcal{X}^2)^M$ is a vector of M QAM modulated signals, $\mathbf{r} \in \mathbb{C}^N$ and $\mathbf{H} \in \mathbb{C}^{N \times M}$ can be written as an equivalent problem in twice the number of real dimensions, i.e., with $\mathbf{r} \in \mathbb{R}^{2N}$ and $\mathbf{H} \in \mathbb{R}^{2N \times 2M}$.



(a) A finite lattice with $M = 2$ and $\mathcal{X} = \{-1, 1\}$.

(b) Some orthogonal projections of a point $\mathbf{y}$.

Fig. 1. Some geometric entities useful in the analysis of sphere decoding.



(a) First expansion, of the target with $i_1 = 1$.

(b) Second expansion, of the level 1 residual in $\mathcal{F}_1(1)$.

Fig. 2. The expansions performed by a sphere decoder in a real 2D search space with BPSK modulation applied to the transmitted signals.

Given matrix $\mathbf{H}$ of full rank M and alphabet $\mathcal{X}$ of size B, we define the *finite lattice* of points in the search set as

$$\mathcal{L} \triangleq \left\{ \mathbf{z} \,\middle|\, \mathbf{z} = \mathbf{H}\mathbf{s},\, \mathbf{s} \in \mathcal{X}^M \right\}. \tag{2}$$

There are B^M lattice points in $\mathcal{L}$, shown in Fig. 1(a) for $B = M = 2$. It can be decomposed into any of M collections of B sub-lattices, each comprised of B^{M-1} lattice points embedded in one of B parallel *affine sets*.[2] Given $i \in \mathcal{I}$, the i^{th} collection

$$\{\mathcal{F}_i(s_i) \mid s_i \in \mathcal{X}\}, \tag{3}$$

contains B affine sets defined as

$$\mathcal{F}_i(s_i) \triangleq \left\{ \mathbf{z} \,\middle|\, \left\langle \mathbf{z} - \mathbf{h}_i s_i, (\mathbf{H}^{-1})_i^T \right\rangle = 0 \right\}, \tag{4}$$

with $\langle \mathbf{a}, \mathbf{b} \rangle$ denoting the inner product. The affine sets in (3) share normal vector $(\mathbf{H}^{-1})_i^T$. We denote the *orthogonal projection* of a vector $\mathbf{y}$ onto this normal as $\text{proj}_{(\mathbf{H}^{-1})_i^T}(\mathbf{y})$; its orthogonal projection onto affine set $\mathcal{F}_i(s_i)$ is then

$$\text{proj}_{\mathcal{F}_i(s_i)}(\mathbf{y}) \triangleq \mathbf{y} - \text{proj}_{(\mathbf{H}^{-1})_i^T}(\mathbf{y} - \mathbf{h}_i s_i), \tag{5}$$

and the corresponding *orthogonal distance* is

$$d(\mathbf{y}, \mathcal{F}_i(s_i)) \triangleq \left| \mathbf{y} - \text{proj}_{\mathcal{F}_i(s_i)}(\mathbf{y}) \right|. \tag{6}$$

Note that $\text{proj}_{\mathcal{F}_i(s_i)}(\mathbf{y})$ is the point in the affine set that is closest in Euclidean distance to $\mathbf{y}$, as depicted in Fig. 1(b).

The B sub-lattices contained in the sets of (3) can be written as shifted copies of one another:

$$\mathcal{L}_i(s_i) = \left\{ \mathbf{z} \,\middle|\, \mathbf{z} = \mathbf{H}_{\setminus i}\mathbf{s}_{\setminus i},\, \mathbf{s}_{\setminus i} \in \mathcal{X}^{M-1} \right\} + \mathbf{h}_i s_i, \tag{7}$$

where we refer to $\mathbf{h}_i$ and s_i as the *offset vector* and *offset coefficient* of $\mathcal{L}_i(s_i)$ and $\mathcal{F}_i(s_i)$, respectively. Observe that $\mathcal{L}_i(s_i) - \mathbf{h}_i s_i$ is a sub-lattice with $M - 1$ degrees of freedom.

Finally, we overview the geometry of sphere decoding. A SD searches for the nearest lattice point to target $\mathbf{r}$ by exploring the affine sets containing sub-lattices of $\mathcal{L}$ in an ordered manner. First it selects index $i_1 \in \mathcal{I}$ and computes the distances from the target to each of the affine sets in the i_1^{th} collection as well as the projections of the target onto each.

$$\mathbf{y}(s_{i_1}) = \text{proj}_{\mathcal{F}_{i_1}(s_{i_1})}(\mathbf{r}),\, s_{i_1} \in \mathcal{X}. \tag{8}$$

The first stage is depicted in Fig. 2(a).

We call projection $\mathbf{y}(s_{i_1})$ a *residual* target, because it represents the components that remain after an orthogonal part has been projected away. Algebraically, it is obtained by applying one of B constraints $\underline{s}_{i_1} = s_{i_1}$. The residual inherits the offset coefficient of the affine set in which it resides, and its *level* is given by the number of constraints applied thusfar in the search. The decomposition of a point into its B residuals is referred to as *expansion*. We use the term *node* to encapsulate the constraints, accumulated distance or cost (from $\mathbf{r}$), level and projection associated with a residual; we also use the terms node and residual interchangeably.

The search problem has now been broken down into B lower dimensional sub-problems, each a search for the nearest sub-lattice point to a residual target:

$$\min_{\mathbf{z} \in \mathcal{L}} |\mathbf{r} - \mathbf{z}|^2 = \min_{s_{i_1} \in \mathcal{X}} \; d^2\left(\mathbf{r}, \mathcal{F}_{i_1}(s_{i_1})\right) + \min_{\mathbf{z} \in \mathcal{L}_{i_1}(s_{i_1})} |\mathbf{y}(s_{i_1}) - \mathbf{z}|^2 \quad . \tag{9}$$

The decoder chooses one of these sub-problems by applying constraint $\underline{s}_{i_1} = s_{i_1}$; all subsequent operations are therefore restricted to affine set $\mathcal{F}_{i_1}(s_{i_1})$. Then it selects another collection $i_2 \in \mathcal{I} \setminus i_1$ and repeats the expansion procedure with residual target $\mathbf{y}(s_{i_1})$ as its argument. Fig. 2(b) illustrates the expansion of a level 1 residual. Observe that one level is gained through each expansion; thus a level M residual is simply a lattice point and its cost is the distance to the target.

In the sequence shown in Fig. 2, we project along $(\mathbf{H}^{-1})_1^T$ in the first expansion, i.e. $i_1 = 1$. This choice imposes a particular order on all of the projections performed by the SD. More formally, we define an *ordering* Π as a permutation of the index set $\mathcal{I}$ and write its elements as the ordered set

$$\Pi = \{i_M, i_{M-1}, \ldots, i_1\}, \tag{10}$$

[2]Recall that an affine set $\mathcal{M} \subset \mathbb{R}^N$ is a set such that $\mathcal{M} = \mathcal{S} + \mathbf{a}$ for some subspace $\mathcal{S} \subset \mathbb{R}^N$ and offset $\mathbf{a} \in \mathbb{R}^N$. See [7, Sec. 1] for more details.

where the mapping from $\mathcal{I}$ to Π is a bijection. Π also defines an $M \times M$ permutation matrix $\mathbf{P}_\Pi \triangleq \left[e_{i_M} \; e_{i_{M-1}} \; \cdots \; e_{i_1} \right]$. We index the elements of Π in reverse order because the relationship between the QR factorization and the F-P or S-E enumerations dictates that a SD selects collections for expansion in the reverse order of the columns of $\mathbf{H}$. For instance, given a 2×2 channel matrix $\mathbf{H}$, to make the decoder project along $(\mathbf{H}^{-1})_1^T$ in the first expansion, we pre-process $\mathbf{H}$ through right multiplication by $\mathbf{P}_{\{2,1\}}$, i.e., by applying ordering $\Pi = \{2, 1\}$.

In the two-dimensional illustration shown in Fig. 2, a SD can find an optimal solution in just two expansions. More generally, in M dimensions or with higher order modulation, this would not necessarily be the case. One key difference between current proposals lies in how the next node for expansion is chosen, both its level and its offset coefficient. For instance, the F-P enumeration expands nodes in order of increasing level, with those at the same level ordered according to their offset coefficients. With respect to Fig. 2(b), it would expand the level 1 residual in $\mathcal{F}_1(-1)$ and then the one in $\mathcal{F}_1(1)$. The S-E enumeration also considers nodes in order of increasing level. However it expands those at the same level in order of increasing orthogonal distance, i.e., in Fig. 2(b), the level 1 residual in $\mathcal{F}_1(1)$ and then the one in $\mathcal{F}_1(-1)$.

Another important consideration is the stop condition for sphere decoding. Most decoders require three input parameters: The received point $\mathbf{r}$, channel matrix $\mathbf{H}$, and a search radius C. In order to guarantee the optimality of the solution returned, the radius must be at least as large as the optimal radius, denoted by C_*, and from a computational perspective, *at least all nodes inside the hypersphere of radius C_* centered at $\mathbf{r}$ must be expanded*. More precisely, we denote the number of nodes expanded by a SD as $\nu(\mathbf{r}, \mathbf{H}, C)$, a quantity that reflects its *computational efficiency* and satisfies

$$\nu(\mathbf{r}, \mathbf{H}, C) \geq \nu(\mathbf{r}, \mathbf{H}, C_*). \tag{11}$$

These ideas are formalized and proved in a separate work [8].

III. Orderings and sphere decoding: An example

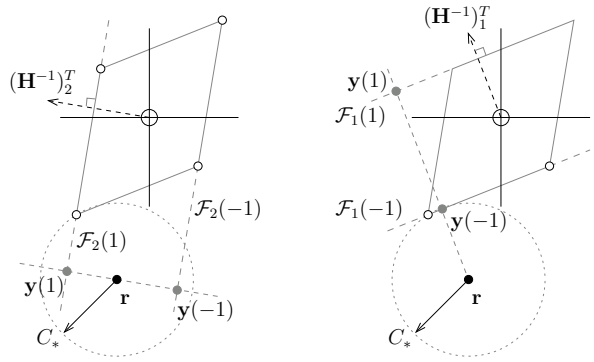
To study the impact of orderings on sphere decoding efficiency, we introduce the notion of a *genie* SD. The genie decoder has prior knowledge of the optimal search radius C_* and expands exclusively those nodes lying within the optimal search hypersphere. It performs the same operations as other SDs, however it is guaranteed to achieve the optimal efficiency

$$\nu_*(\mathbf{r}, \mathbf{H}) \triangleq \nu(\mathbf{r}, \mathbf{H}, C_*). \tag{12}$$

Unlike other decoders, when the genie decoder is called with different orderings applied to the channel matrix, any variations in ν_* can be attributed solely to the ordering, and not to other parameters that may be more difficult to capture, e.g., the choice of the next residual for expansion. This approach enables us to effectively decouple the ordering part of the problem from the SD itself.

Next we consider how the genie decoder would decode a received signal transmitted using BPSK modulation over a sample 2×2 channel

$$\mathbf{H} = \begin{bmatrix} 1.13 & -5.65 \\ 6.78 & -2.20 \end{bmatrix}. \tag{13}$$



(a) Natural ordering: Decoder projects along $(\mathbf{H}^{-1})_2^T$ first; three expansions are needed.

(b) Inverted ordering: Decoder projects along $(\mathbf{H}^{-1})_1^T$ first; two expansions are needed.

Fig. 3. The expansions performed when decoding $\mathbf{r}$ (transmitted using BPSK modulation) under the two possible orderings of channel matrix $\mathbf{H}$.

Given a received point $\mathbf{r}$ and ordering $\Pi = \{1, 2\}$, i.e., the natural ordering, the decoder begins by expanding $\mathbf{r}$ along normal vector $(\mathbf{H}^{-1})_2^T$. Fig. 3(a) shows the resulting expansions; in particular, note that there are three nodes inside the detection hypersphere and so $\nu_*(\mathbf{r}, \mathbf{H}[\,\mathbf{e}_1 \; \mathbf{e}_2\,]) = 3$.

The neighbouring plot in Fig. 3(b) shows the behaviour of the genie decoder under the inverted ordering $\Pi = \{2, 1\}$. In this case, the first expansion is along normal vector $(\mathbf{H}^{-1})_1^T$ and the resulting efficiency is $\nu_*(\mathbf{r}, \mathbf{H}[\,\mathbf{e}_2 \; \mathbf{e}_1\,]) = 2$. Therefore we might conclude that the inverted ordering is advantageous for improved decoder efficiency, given the received point in question. It is also important to observe that ν_* is governed by the geometric locations of the nodes, which are in turn determined solely by $\mathbf{r}$ and the ordering applied to $\mathbf{H}$.

More generally, Fig. 4 shows a map of the optimal ordering regions of the columns of sample channel matrix $\mathbf{H}$ over the domain of possible received points $\mathbf{r}$, where we call ordering Π_* and permuted channel matrix $\mathbf{H}_* = \mathbf{H}\mathbf{P}_{\Pi_*}$ *optimal* if

$$\nu_*(\mathbf{r}, \mathbf{H}_*) \leq \nu_*(\mathbf{r}, \mathbf{H}\mathbf{P}) \tag{14}$$

for all $M!$ permutation matrices $\mathbf{P}$.

Given our sample channel matrix, $M! = 2$ and there are two possible orderings of $\mathbf{H}$. In Fig. 4, light shading indicates points $\mathbf{r}$ for which $\nu_*(\mathbf{r}, \mathbf{H}[\,\mathbf{e}_1 \; \mathbf{e}_2\,]) = \nu_*(\mathbf{r}, \mathbf{H}[\,\mathbf{e}_2 \; \mathbf{e}_1\,])$, i.e., where both orderings result in the same decoder efficiency, medium shading shows where the decoder using ordering $\Pi = \{1, 2\}$ expands fewer nodes, and dark shading, where ordering $\Pi = \{2, 1\}$ is favoured. As expected, since the point considered in Fig. 3 is located in a darkly shaded region, the inverted ordering leads to improved decoding efficiency.

More importantly, Fig. 4 establishes an important property: *The optimal ordering for efficient sphere decoding depends not*

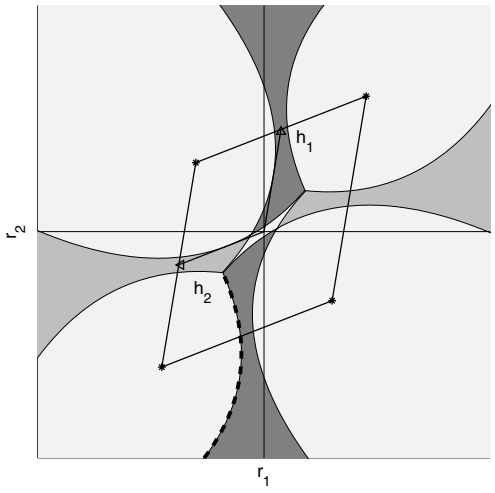


Fig. 4. Optimal ordering of the columns of channel matrix $\mathbf{H}$ as a function of received point $\mathbf{r}$: Light shading indicates that both orderings are equivalent, medium shading that $\mathbf{H}_* = \mathbf{H}[\,\mathbf{e}_1\ \mathbf{e}_2\,]$, and dark shading that $\mathbf{H}_* = \mathbf{H}[\,\mathbf{e}_2\ \mathbf{e}_1\,]$. The dashed curve is highlighted for further analysis.

only on the channel matrix $\mathbf{H}$*, but also on the received point* $\mathbf{r}$. It also provides insight into why some orderings achieve better efficiencies than others for a given received point $\mathbf{r}$. Consider the dashed curve highlighted in Fig. 4. It delineates the boundary between a region where $\mathbf{H}_* = \mathbf{H}[\,\mathbf{e}_2\ \mathbf{e}_1\,]$ and one where both orderings result in the same efficiency. The change arises because the efficiency of the natural ordering is reduced as the boundary is crossed from left to right.

To understand this behaviour, consider the projections performed by the genie decoder under ordering $\{1, 2\}$, as $\mathbf{r}$ crosses the boundary. An equation for this curve is given by

$$d(\mathbf{r}, \mathcal{F}_2(-1)) = C_*, \tag{15}$$

defined over regions where $\mathbf{r}$ is such that the optimal search radius satisfies $C_* = \left|\mathbf{r} - \mathbf{H}\begin{bmatrix}-1\\1\end{bmatrix}\right|$.

When $\mathbf{r}$ lies to the left of the boundary, the distance from $\mathbf{r}$ to the *second nearest* affine set in the chosen collection satisfies $d(\mathbf{r}, \mathcal{F}_2(-1)) > C_*$, and so the decoder does not have to expand *any* of the nodes contained in $\mathcal{F}_2(-1)$. However, when $\mathbf{r}$ crosses to its right, $d(\mathbf{r}, \mathcal{F}_2(-1)) < C_*$ and consequently node $\mathbf{y}(-1)$ lies within the optimal search hypersphere. Since all nodes lying within this hypersphere must be expanded, the efficiency of the natural ordering becomes degraded by one.

IV. A new ordering scheme

Now we formalize the intuition gained from the previous example, generalize it to higher dimensions, and present the new ordering. Our overall strategy is, at each level, to choose the best collection of affine sets onto which to project the residual target $\mathbf{y}$. The following decision criterion is proposed:

$$i' = \underset{i}{\operatorname{argmax}}\, d(\mathbf{y}, \mathcal{F}_i(\beta_i)), \tag{16}$$

where β_i is the offset coefficient of the second nearest affine set to $\mathbf{y}$ in collection i.

The rationale behind decision criterion (16) is to expand $\mathbf{y}$ along the direction where the second nearest affine set is the greatest distance away. In so doing, we seek to reduce the number of nodes expanded by the genie decoder in *any* of the affine sets $\mathcal{F}_{i'}(s_{i'})$, $s_{i'} \neq \alpha_{i'}$, where $\alpha_{i'}$ is the offset coefficient of the *nearest* affine set to $\mathbf{y}$ in the chosen collection. In the best case, if all of the expanded nodes lie in $\mathcal{F}_{i'}(\alpha_{i'})$, the problem dimension will effectively be reduced by one, leading to a greatly improved decoder efficiency.

Having chosen collection i', we propose to project $\mathbf{y}$ onto $\mathcal{F}_{i'}(\alpha_{i'})$. The remaining normal vectors $(\mathbf{H}^{-1})_i^T$, $i \neq i'$ are likewise projected in order to restrict subsequent operations to the same affine set. The decision metrics can then be recomputed in the lower dimensional set and this procedure repeated until all M selections have been made. Pseudocode to compute proposed ordering $\widehat{\Pi}$ is given in Algorithm 1:

Algorithm 1 An Enhanced Ordering $(\mathbf{r}, \mathbf{H}, M, \mathcal{X})$

$\mathcal{I} \longleftarrow \{1, 2, \ldots, M\}$ Initialize index set
$\mathbf{y} \longleftarrow \mathbf{r}$ Initialize target
$\mathbf{G} \longleftarrow (\mathbf{H}^{-1})^T$ Compute inverse
for each level L from 1 to M **do**
  for each index i in set $\mathcal{I}$ **do**
    $\alpha_i \longleftarrow \operatorname{argmin}_{x\in\mathcal{X}} |\langle \mathbf{y}, \mathbf{g}_i\rangle - x|$ Find nearest sets
    $\beta_i \longleftarrow \operatorname{argmin}_{x\in\mathcal{X}\setminus\alpha_i} |\langle \mathbf{y}, \mathbf{g}_i\rangle - x|$ Second nearest
    $\delta_i \longleftarrow d\left(\mathbf{y}, \mathcal{F}_i(\beta_i)\right)$ Compute distances
  end for
  $\widehat{\Pi}_{M-L+1} \longleftarrow i' \longleftarrow \operatorname{argmax}_{i\in\mathcal{I}} \delta_i$ Select collection L
  $\mathbf{s}_{0,i'} \longleftarrow \alpha_{i'}$ Record offset coeff.
  $\mathcal{I} \longleftarrow \mathcal{I} \setminus i'$ Remove from index set
  $\mathbf{y} \longleftarrow \operatorname{proj}_{\mathcal{F}_{i'}(\alpha_{i'})}(\mathbf{y}) - \mathbf{h}_{i'}\alpha_{i'}$ Project+shift residual
  for each index i in set $\mathcal{I}$ **do** Project inverse
    $\mathbf{g}_i \longleftarrow \operatorname{proj}_{\mathcal{F}_{i'}(0)}(\mathbf{g}_i)$
  end for
end for
Return $\widehat{\Pi}$, $\mathbf{s}_0$

Recall that projecting onto the affine set with offset coefficient $\alpha_{i'}$ in collection i' is algebraically equivalent to applying constraint $s_{i'} = \alpha_{i'}$. Recording these constraints as the ordering is computed provides an initial data estimate $\mathbf{s}_0$ as a byproduct of Algorithm 1. Further results on the properties of $\mathbf{s}_0$ are discussed in a separate work [9].

V. Performance Evaluation

Here we provide two ways of evaluating the performance of the enhanced ordering scheme. First we consider how closely the orderings recommended by Algorithm 1 correspond to the optimal ones confirmed by simulation (Fig. 4). Fig. 5 depicts a map of the proposed ordering decisions $\widehat{\mathbf{H}} = \mathbf{H}\mathbf{P}_{\widehat{\Pi}}$ over the domain of possible received points $\mathbf{r}$, given the same sample channel as before (13). For this two-dimensional example, the match between the recommendations made by Algorithm 1 and the optimal orderings can easily be verified graphically. A similar correspondence is evidenced when using higher order modulations; these results have been omitted for brevity.

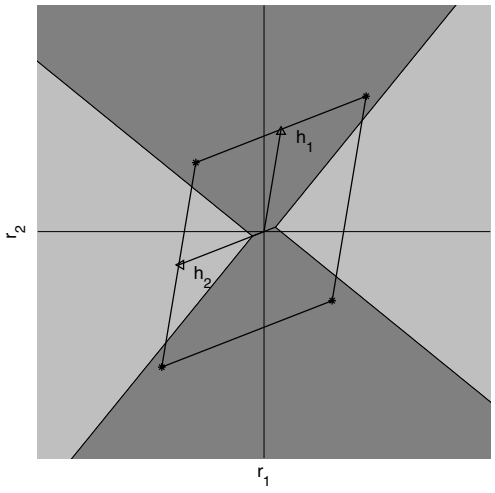


Fig. 5. Proposed ordering of the columns of a sample two-dimensional (real) channel matrix $\mathbf{H}$ as a function of received point $\mathbf{r}$: Medium shading indicates where $\mathbf{H}_\mathrm{p} = \mathbf{H}[\,\mathbf{e}_1\ \mathbf{e}_2\,]$, and dark shading where $\mathbf{H}_\mathrm{p} = \mathbf{H}[\,\mathbf{e}_2\ \mathbf{e}_1\,]$.

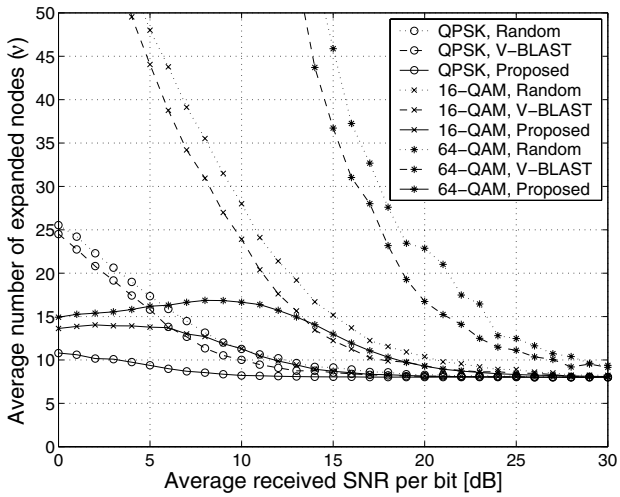


Fig. 6. Average number of nodes expanded by a SD based on the S-E enumeration vs. average received SNR per bit for the random, V-BLAST and proposed orderings over a 4:4 MIMO (complex) flat fading channel.

One of the key hypotheses underlying our work is that an ordering effective in enhancing the performance of the genie decoder should also enhance that of other SDs based on the F-P or S-E enumerations. Thus a more important performance evaluation of the enhanced ordering scheme is obtained by considering how successful it is in reducing the number of nodes expanded by a standard SD. Fig. 6 shows the average number of nodes expanded by an efficient decoder based on the S-E enumeration that adaptively decreases its search radius as lattice points are discovered [2]. The performance curves obtained under three orderings are reported: Random ordering, that used in V-BLAST decoding [10], and the new proposal.

A vast improvement is realized by the new ordering, especially at low SNRs, where the complexity of existing SDs is not widely considered to be competitive. What is particularly noteworthy is that this improvement remains significant even as the modulation order is increased. Although the ordering was derived using the notion of a genie decoder, simulation results indicate that the benefits offered are transferable to other SDs as well. Thus we have shown that the proposed scheme is an extremely effective pre-processing stage that can be readily combined with existing algorithms.

VI. Discussion and Conclusions

In this paper we have proposed a novel yet compatible pre-processing stage for sphere decoders. We demonstrated via simulation that it not only offers a significant reduction in the computational cost of subsequent sphere decoding stages, but also that it is especially effective in the low SNR regime, which has traditionally been an expensive one for SDs. Another strength of the new ordering is that it is of increasing influence at higher spectral efficiencies, thus enabling existing decoders to achieve ML detection at a competitive computational cost over a wider range of performance parameters.

The time complexity of Algorithm 1 is $\mathcal{O}(M^3)$, roughly comparable to a matrix inversion and a few QR factorizations. In practice, we have observed immensely reduced decoding times at low SNRs using unoptimized implementations of the proposed ordering. As can be seen from Fig. 6, as the SNR exceeds a certain modulation-dependent level, the number of nodes expanded by most SDs, even under random ordering, converges to a lower bound.[3] Thus at such SNRs it may not be efficient to apply any pre-processing.

References

[1] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639–1642, July 1999.

[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, October 2003.

[3] A. M. Chan and I. Lee, "A new reduced-complexity sphere decoder for multiple antenna systems," in *IEEE International Conference on Communications*, vol. 1, April 2002, pp. 460–464.

[4] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm: Part I, The expected complexity," *To appear in IEEE Transactions on Signal Processing*, 2004.

[5] D. Seethaler, H. Artés, and F. Hlawatsch, "Dynamic nulling-and-cancelling with near-ML performance for MIMO communication systems," in *IEEE Int'l Conf. on Acoustics Speech and Signal Processing*, vol. 4, May 2004, pp. 777–780.

[6] A. Wiesel, X. Mestre, A. Pagés, and J. R. Fonollosa, "Efficient implementation of sphere demodulation," in *IEEE Workshop on Signal Processing Advances in Wireless Communications*, June 2003.

[7] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 1970.

[8] K. Su, C. N. Jones, and I. J. Wassell, "An automatic sphere decoder," *Submitted to IEEE Transactions on Information Theory*, 2004.

[9] K. Su and I. J. Wassell, "Efficient MIMO detection by successive projection," *Submitted to IEEE Int'l Symp. on Information Theory*, 2005.

[10] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *International Symposium on Signals, Systems, and Electronics*, September 1998, pp. 295–300.

[3] For all SDs, there is a lower bound of $\nu \geq 2M$, where we recall that M is the number of transmit antennas and the factor of two arises because there are two real dimensions per complex dimension.