A Tree Pruning Algorithm For MIMO Sphere Decoding Based On Path Metric

Shiliang Wang*, Xiaolong Guo*, Songlin Sun*, Tiehong Tian†, Shizhen Sun†, Xiaojun Jing*
*Key Laboratory of Trustworthy Distributed Computing and Service (BUPT), Ministry of Education
Beijing University of Posts and Telecommunications, Beijing, China, 100876
Email: wangslzxr@gmail.com
†China Unicom System Integration Limited Corporation

***Abstract*—Tree pruning can significantly reduce the complexity of sphere decoding (SD). How to determine the pruning rule is an open problem of tree pruning. In this paper, we propose a pruning strategy for SD based on path metric. Because only the nearest lattice point is concerned, if the ratio of the metric to the minimum metric is larger than a threshold, the path whose metric is large enough can be pruned. We analyze the influence of the choice of the thresholds on the performance and the complexity. Through analysis and the simulations, we can show that the complexity reduction is significant while maintaining the negligible performance degradation when proper thresholds are chosen. Besides, tradeoff between complexity and performance can be easily achieved by adjusting the thresholds.**

I. Introduction

Multiple-input multiple-output (MIMO) communication systems have been shown to be capable of achieving huge capacity on a scattering-rich wireless channel [1], [2]. Maximum likelihood decoding (MLD) is known as the optimum detection method, but its exponential complexity with the number of transmit antennas and the constellation size, because of the requirement of exhaustive full search, makes it unrealizable in a practical system. Sphere decoding (SD) is a method to reduce the complexity of MLD [3], [4].

Even though the SD algorithm offers computational efficiency in many communication scenarios [5], [6], it also requires exponential complexity [7] both from a worst-case and from an average point of view. Many algorithms have been proposed to further reduce the complexity of SD. One of them is radius control which is to adjust the search process after obtaining a candidate point or to choose a proper initial radius since the radius determines the search space and influences the complexity of SD significantly, such as increasing radius search (IRS) [7], inter-searching radius control (ISRC) [8], stopping radius for the sphere decoder (SRSD) [9] and so on. Tree pruning, which sets different radius for different layers, is another method. It can remove the unlikely branches in early stage of sphere search, so that the number of visited nodes is reduced, such as increasing radii algorithm (IRA) [10] and probabilistic tree pruning sphere decoding (PTP-SD) [11] algorithm. It has been shown that these algorithms reduce the complexity at the expense of negligible performance loss.

This research is financially supported by the National Natural Science Foundation of China (NSFC) (Grant No. 60702049 and 60872148)

Motivated by the fact that the sphere constraint of the SD algorithm offers a loose necessary condition, especially in the early layers of search, tree pruning methods further reduce the computational complexity by exploiting an aggressive pruning strategy, which is realized by setting different sphere radius per layer. IRA, proposed by Gowaikar and Hassibi, analyzes the relation among the radii, the performance and the complexity. Based on the mean of the path metrics, IRA settles upon a linear schedule for the radii, and the slope of the linearity is the variance of the noise. PTP-SD chooses the radii based on a probability model, which can obtain the probability that the current path will correspond to the correct point by adding the probabilistic noise constraint into the path metrics. These strategies are based on the noise statistics, and the estimation of signal-to-noise ratio (SNR) is necessary.

Our strategy, named PMP-SD, makes tree pruning according to the path metrics, without estimation of SNR. For MLD, only the closest lattice point with the minimum cost metric is concerned. So the larger the path metric of a path is, the lower the probability that the path corresponds to the correct point is. The proposed algorithm is to store the current minimum path metric for each layer and to compare current path metric to the minimum. The path will be pruned as long as the ratio of the two path metrics, which indicates the difference between them, is greater than the threshold of this layer. Besides, we analyze how to choose the thresholds and discuss its influence on performance and complexity. Through analysis and simulations, we can show that the complexity reduction is significant while maintaining the negligible performance degradation and the tradeoff between complexity and performance can be easily achieved by adjusting the thresholds.

The remainder of this paper is organized as follows. In section II, we will present the system model and review the classic sphere decoding algorithm with Schnorr-Euchner enumeration [4]. We will present our algorithm in section III. The simulation results will be provided in section IV and section V will be the conclusion.

II. Sphere Decoding

A. System Model

Considering the uncoded MIMO system of M transmit and N receive antennas for $M \geq N$, the received signal at each

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

instant time is given by

$$\mathbf{y}_c = \frac{1}{\sqrt{ME}}\mathbf{H}_c\mathbf{x}_c + \mathbf{v}_c \tag{1}$$

where $\mathbf{x}_c$ is the transmitted symbol vector whose components are elements of a quadrature amplitude modulation (QAM) signal set $\mathcal{A}$ with size A. $\mathbf{y}_c$ is the complex received signal vector. $\mathbf{H}_c = [h'_{ij}]$ is a complex channel matrix without any particular structure, known perfectly to the receiver. We assume $h'_{ij} \sim \mathcal{CN}(0,1)$, and $\mathbf{v}_c$ is the complex noise vector whose components are independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian random variables. E is the average power of transmitted symbol. Therefore, the average power of the received signal at each receive antenna is 1. If the SNR is ρ, then the variance of the component of $\mathbf{v}_c$ is $1/\rho$.

The maximum likelihood (ML) decoding is given by

$$\hat{\mathbf{x}}_c = \arg\min_{\mathbf{x}_c \in \mathcal{A}^M} |\mathbf{y}_c - \frac{1}{\sqrt{ME}}\mathbf{H}_c\mathbf{x}_c|^2 \tag{2}$$

In order to use SD, the complex number signal model in (1) needs to be reformulated to a real number signal model,

$$\mathbf{y} = \begin{bmatrix}\Re(\mathbf{y}_c)\\ \Im(\mathbf{y}_c)\end{bmatrix}, \mathbf{x} = \begin{bmatrix}\Re(\mathbf{x}_c)\\ \Im(\mathbf{x}_c)\end{bmatrix}, \mathbf{v} = \begin{bmatrix}\Re(\mathbf{v}_c)\\ \Im(\mathbf{v}_c)\end{bmatrix}$$

$$\mathbf{H} = \frac{1}{\sqrt{ME}}\begin{bmatrix}\Re(\mathbf{H}_c) & -\Im(\mathbf{H}_c)\\ \Im(\mathbf{H}_c) & \Re(\mathbf{H}_c)\end{bmatrix} \tag{3}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ are the real and imaginary parts of its argument. Then the real number signal model is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{4}$$

Let $m = 2M$, $n = 2N$, so $\mathbf{H}$ is a $n \times m$ real matrix, where h_{ij} is an entry of $\mathbf{H}$. Its mean and variance are zero and $\sigma_h^2 = 1/2ME$ respectively. The entries of $\mathbf{x}$, say x_i, belong to a pulse amplitude modulation (PAM) signal set $\mathcal{B}$, whose size is $s = \sqrt{A}$, i.e.

$$\mathcal{B} = \{x = 2b - s + 1 | b \in \mathcal{Z}_s\}$$

where $\mathcal{Z}_s = \{0, 1, \ldots, s-1\}$. $\mathbf{v} \in \mathcal{R}^n$ and its entries, denoted as v_i, are i.i.d. $\mathcal{N}(0, \sigma^2)$ random variables, where $\sigma^2 = 1/2\rho$ is the power of real noise. Then the real MLD is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{B}^m} |\mathbf{y} - \mathbf{H}\mathbf{x}|^2 \tag{5}$$

This is a finite-alphabet-constrained least-squares (LS) problem, which is known as nondeterministic polynomial-time (NP)-hard. Its complexity is exponential with the number of transmit antenna and the constellation size.

B. Sphere Decoding

Usually, (5) is considered as a *closest lattice point search* problem. The set $\Lambda = \{\mathbf{H}\mathbf{x} | \mathbf{x} \in \mathcal{B}^m\}$ is a finite lattice. The search in (5) is to obtain the lattice point which is the closest to the given point $\mathbf{y}$. This problem has been widely investigated in lattice theory.

SD reduces the complexity by limiting the search space in a hyper sphere $S(\mathbf{y}, \sqrt{C})$ centered at $\mathbf{y}$, where C is the squared radius of the sphere. SD can be expressed as

$$|\mathbf{y} - \mathbf{H}\mathbf{x}|^2 \leq C \tag{6}$$

Performing QR-decomposition of $\mathbf{H}$ as $\mathbf{H} = \begin{bmatrix}\mathbf{Q} & \mathbf{Q}'\end{bmatrix}\begin{bmatrix}\mathbf{R}^T & \mathbf{0}^T\end{bmatrix}^T$, where $\mathbf{R}$ is an $m \times m$ upper triangular matrix with positive diagonal entries, $\mathbf{0}$ is a zero matrix, and $\mathbf{Q}$ (resp., $\mathbf{Q}'$) is an $n \times m$ (resp., $n \times (n-m)$) unitary matrix. The inequality (6) is equivalent to

$$\begin{aligned}\left|\begin{bmatrix}\mathbf{Q} & \mathbf{Q}'\end{bmatrix}^T \mathbf{y} - \begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}\mathbf{x}\right|^2 &\leq C\\ \left|\mathbf{Q}^T\mathbf{y} - \mathbf{R}\mathbf{x}\right|^2 + \left|(\mathbf{Q}')^T\mathbf{y}\right|^2 &\leq C\\ |\mathbf{y}' - \mathbf{R}\mathbf{x}|^2 &\leq c_0\end{aligned}$$

where $\mathbf{y}' = \mathbf{Q}^T\mathbf{y}$, $c_0 = C - c_s$ and $c_s = |(\mathbf{Q}')^T\mathbf{y}|^2$. Obviously, c_s is known to the receiver and independent of $\mathbf{x}$. Then the above inequality can be written as

$$\sum_{i=1}^{m}\left(y'_i - \sum_{j=i}^{m} r_{ij}x_j\right)^2 = \sum_{i=1}^{m} B_i \leq c_0 \tag{7}$$

where r_{ij} denotes an (i,j) entry of $\mathbf{R}$, and

$$B_i = \begin{cases}\left(y'_i - \sum_{j=i}^{m} r_{ij}x_j\right)^2 & 1 \leq i \leq m\\ \left(\sum_{j=1}^{n} q'_{j,i-m}y_j\right)^2 & m+1 \leq i \leq n\end{cases}$$

Because B_i is known to the receiver for $m+1 \leq i \leq n$ and independent of $\mathbf{x}$, we only consider the B_i for $1 \leq i \leq m$. We can find that B_i depends only on $x_m, \ldots, x_i$ in this case. By considering the conditions in order from m to 1, we obtain the range of x_i for given values of symbols $x_m, \ldots, x_{i+1}$. If the partial vector $\mathbf{x}_{i+1}^m = [x_{i+1}, \ldots, x_m]^T$ is in the sphere, the range of x_i is not a null set. The lower bound of x_i is

$$L_i = \left\lceil \frac{1}{r_{ii}}\left(y'_i - \sum_{j=i+1}^{m} r_{ij}x_j - \sqrt{c_0 - \sum_{j=i+1}^{m} B_j}\right)\right\rceil \tag{8}$$

and the upper bound is

$$U_i = \left\lfloor \frac{1}{r_{ii}}\left(y'_i - \sum_{j=i+1}^{m} r_{ij}x_j + \sqrt{c_0 - \sum_{j=i+1}^{m} B_j}\right)\right\rfloor \tag{9}$$

Otherwise the range is a null set and x_{i+1} should take the next admissible value.

Obviously, SD can be considered as a decision tree with $m+1$ layer, and s branches emanating from each node. We define B_i as the metric of the branch. The partial vector $\mathbf{x}_i^m = [x_i, \ldots, x_m]^T$ determines a path traveling from the root to the node x_i passing $x_m, \ldots, x_{i+1}$. The cumulative metric $P_i(\mathbf{x}_i^m) = \sum_{k=m+1}^{n} B_k + \sum_{k=i}^{m} B_k(\mathbf{x}_k^m)$ is defined as path metric, P_i for short. Because $\sum_{k=m+1}^{n} B_k = c_s$ is known to the receiver, we define it as the path metric of the root node of the decision tree. When SD reaches the bottom layer, a candidate point $\tilde{\mathbf{x}}$ is obtained. The squared Euclidean distance from the lattice point to the center $\mathbf{y}$, say cost metric, is denoted as $D(\tilde{\mathbf{x}})$, i.e. $D(\tilde{\mathbf{x}}) = |\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}|^2$. Note $D(\tilde{\mathbf{x}}) = P_1(\tilde{\mathbf{x}})$.

There are two strategies to enumerate all the points. Searching order in each layer of Pohst enumeration [3] is the natural spanning. The other strategy called Schnorr-Euchner enumeration (SE) [4] where the points are sorted and examined basing on their branch metric values is more attractive due to its low complexity [6]. For a detailed discussion of the SE algorithm, the reader is referred to Algorithm II proposed in [6]. In this paper, the proposed algorithm is based on SE, though it can also be used for Pohst enumeration.

III. Statistic Tree Pruning Algorithm Based On Path Metric

A. PMP-SD

In MLD, only the lattice point with the minimum cost metric, denoted as $\mathbf{x}_{ML}$ is desired. Although it is not necessary that all the path metrics of $\mathbf{x}_{ML}$ are also minimum, the path is unlikely to correspond to the $\mathbf{x}_{ML}$ if its path metric is much larger than the minimum path metric at the same layer. Therefore, the path can be early pruned.

The idea of PMP-SD is based on the fact mentioned above. In PMP-SD, we use another m registers to store the current minimum path metrics for each layer, denoted as $\overline{\mathbf{P}} = \begin{bmatrix}\overline{P}_1 & \overline{P}_2 & \dots & \overline{P}_m\end{bmatrix}$. As the research process continues, the stored minimum path metrics will be updated when smaller metrics are obtained. However, if the metric of a path is much larger than the corresponding stored minimum, the path will be pruned.

We define the ratio of current path metric to the minimum metric at the same layer as the indicator in order to determine if current path metric is large enough to be pruned, i.e.

$$\gamma = \frac{P_i\left(\mathbf{x}_i^m\right)}{\overline{P}_i}$$

If γ is larger than a predefined threshold, the path metric is large enough and the path is unlikely to correspond to the $\mathbf{x}_{ML}$, so the path will be pruned. Otherwise, the search will continue. Because the property of the path metric ratio is different, the ratio thresholds for each layer are different, too. We save the thresholds in a vector $\overline{\boldsymbol{\gamma}} = \begin{bmatrix}\overline{\gamma}_1 & \overline{\gamma}_2 & \dots & \overline{\gamma}_m\end{bmatrix}$. The pruning rule can also be expressed as path metric threshold, i.e. if

$$P_i\left(\mathbf{x}_i^m\right) > \overline{\gamma}_i \overline{P}_i$$

the path will be pruned. Or, the search will continue.

The conventional statistical tree pruning algorithm determines the pruning rule based on the noise statistics. Therefore, dynamic and accurate SNR value measurement in order to achieve optimal performance is necessary. In the case of wrong estimation of SNR, the performance can be degraded and the complexity can be increased. PMP-SD is SNR measurement-free, thus the thresholds of path metric can be adaptively adjusted according to the noise and the channel matrix since the minima are dynamically changed and the ratio thresholds are fixed.

In order to achieve the near ML performance and complexity saving, the thresholds need to be chosen deliberately. Generally, too small thresholds result in the performance loss and too large thresholds will not be helpful for reducing the complexity. Since the lattice point with minimum cost metric is chosen, the ratio threshold at layer 1, i.e. $\overline{\gamma}_1$, is always set to 1. The choice of the thresholds for the other layers and the influence on the performance and complexity is discussed at III.C. The algorithm based on SE enumeration is formalized as follows.

PMP-SD (**Input** $\mathbf{y}'$, $\mathbf{R}$, c_s, C and $\overline{\boldsymbol{\gamma}}$. **Output** $\hat{\mathbf{x}}$)

Step 1. Set $i := m$, $B_m := 0$, $P_{m+1} := c_s$, $\xi_m := 0$, $d_c := C$ and $\overline{P}_k := \infty$ for $k = 1, \dots, m$.

Step 2. Set $x_i := \lfloor (y_i' - \xi_i)/r_{ii} \rceil$ and $\Delta_i :=$sign$(y_i' - \xi_i - r_{ii}x_i)$.

Step 3. Set $B_i := (y_i' - \xi_i - r_{ii}x_i)^2$ and $P_i := P_{i+1} + B_i$. If $d_c < P_i$, then go to step 4. Else if $x_i \notin$ Constellation, go to step 6. Else if $i > 1$, go to step 7. Else($i = 1$), go to step 5.

Step 4. If $i = m$, terminate. Else set $i := i + 1$ and go to step 6.

Step 5. Let $d_c := P_1$, save $\hat{\mathbf{x}} := \mathbf{x}$. Set $i := i+1$ and go to step 6.

Step 6. Let $x_i := x_i + \Delta_i$, $\Delta_i := -\Delta_i -$sign(Δ_i), and go to step 3.

Step 7. If $P_i \geq \overline{\gamma_i}\overline{P_i}$, then go to step 4. Else if $P_i < \overline{P}_i$, then $\overline{P}_i := P_i$. Set $\xi_{i-1} := \sum_{j=i}^{m} r_{i-1,j}x_j, i := i-1$ and go to step 2.

B. Preprocessing

Because the pruning rule of PMP-SD is based on path metric, so the performance of PMP-SD is highly dependent on the values of diagonal entries of $\mathbf{R}$. The larger the absolute values of diagonal entries of $\mathbf{R}$, the larger the SNR.

Because the detection is started from the right bottom of $\mathbf{R}$, we should sort the diagonal entries of $\mathbf{R}$ to make the right bottom part larger than that of the left top part. We propose the heuristic sorted QR-decomposition algorithm proposed in [12] due to its good balance between performance and complexity.

C. Performance and The Choice of Ratio Thresholds

According to the algorithm, if the path metric of the transmitted vector $P_i\left(\mathbf{x}_i^m\right)$ is larger than $\overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)$ and $\mathbf{x} \neq \tilde{\mathbf{x}}$, an error occurs. Therefore, for all $\tilde{\mathbf{x}}_i^m \in \mathcal{B}^l$, where $l = m-i+1$, $P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)$, the decoding is correct, i.e. the vector error probability is

$$\begin{aligned}
&Pr\left(e\right) \\
&= 1 - \prod_{i=1}^{m} \prod_{\tilde{\mathbf{x}}_i^m \in \mathcal{B}^l} Pr\left[P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right] \\
&= 1 - \left[1 - P_{ML}^e\right] \prod_{i=2}^{m} \prod_{\tilde{\mathbf{x}}_i^m \in \mathcal{B}^l} Pr\left[P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right] \\
&\leq P_{ML}^e + \epsilon
\end{aligned}$$

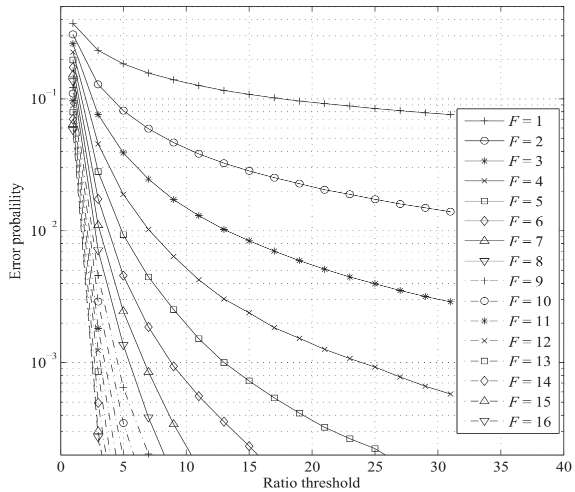

Fig. 1. $Pr\left[P_i\left(\mathbf{x}_i^m\right) > \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right]$ for different degrees of freedom from 1 to 16

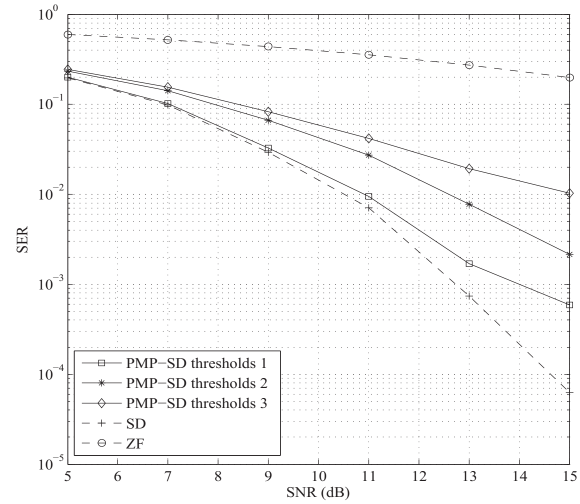

Fig. 2. Performance of PMP-SD for 8×8 MIMO system with 4-QAM

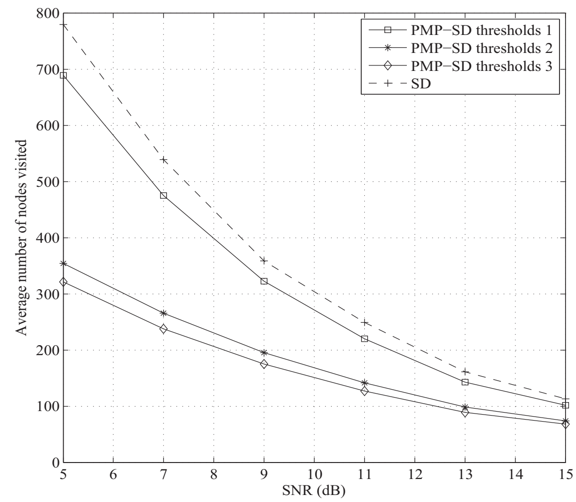

Fig. 3. Complexity of PMP-SD for 8×8 MIMO system with 4-QAM

where P_{ML}^e is the error probability of MLD, $\epsilon = 1 - \prod_{i=2}^{m} \prod_{\tilde{\mathbf{x}}_i^m \in \mathcal{B}^l} Pr\left[P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right]$ is an additional error to MLD and it is a decreasing function on $\overline{\gamma}_i$s.

When decoding is perfect, the branch metric of $\mathbf{x}_i^m$ only contains noise, so $D(\mathbf{x}) = |\mathbf{v}|^2$ is a scaled chi-squared random variable with n degrees of freedom and its probability density function (PDF) is $\Gamma\left(n/2, 1/2\sigma^2\right)$, where $\Gamma\left(\cdot\right)$ is the Gamma function. Additionally, path metric $P_i\left(\mathbf{x}_i^m\right)$ is also scaled chi-squared random variable with $n - i + 1$ degrees of freedom. Its PDF is $\Gamma\left((n-i+1)/2, 1/2\left(\sigma^2 + \sigma_h^2\left|\mathbf{e}_i^m\right|^2\right)\right)$ [7], where $\mathbf{e}_i^m = \mathbf{x}_i^m - \tilde{\mathbf{x}}_i^m$ is the error vector. Therefore, $Pr\left[P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right]$ is the probability that a scaled chi-squared random variable is smaller than the other one with a larger variance which depends on the error vector. [7] gives a method to calculate $\prod_{\tilde{\mathbf{x}}_i^m \in \mathcal{B}^l} Pr\left[P_i\left(\mathbf{x}_i^m\right) \leq \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)\right]$ by counting the number for each possible value of $\left|\mathbf{e}_i^m\right|^2$. For example, there are $\binom{l}{i/4}$ vectors whose $\left|\mathbf{e}_i^m\right|^2$ equals i when 4-QAM is used at the transmit antennas. But it is also difficult to determine the ratio thresholds by the conclusion because it is too complex.

In order to get more insight into the choice strategy, we make a simplification. Generally, the most dominant error is the case where only one entry of $\tilde{\mathbf{x}}$ is not equal to its corresponding entry in $\mathbf{x}$. Therefore, we can only consider this case. Fig. 1 describes the error probability $P_i\left(\mathbf{x}_i^m\right) > \overline{\gamma}_i P_i\left(\tilde{\mathbf{x}}_i^m\right)$ as a function of the ratio thresholds for 8×8 MIMO with 4-QAM at all the degrees of freedom, denoted as F in the figure. The SNR is set to be 10dB. We can see that the error probability will be almost always large at the early layers where Fs are small even when ratio thresholds are very large. For example, the error probability cannot be smaller than 0.05 even when $\overline{\gamma}_1 = 32$. But when Fs are large, the probability decreased quickly and the ratios are very similar for a specific probability. Although the figure just describe the situation for SNR=10dB, they are similar at low and medium SNR regime where the complexity of SD is very high. From this, we can determine a ratio thresholds choice strategy. At the top layers, the thresholds should be large or even without pruning at the first layer since it is very dangerous to prune any path. At the bottom, the thresholds can be very small.

IV. SIMULATION

In this section, we will compare the performance and the complexity of PMP-SD, MLD, PTP-SD and zero-forcing (ZF) decoder. Additionally, the influence of the sorted QR-decomposition is also simulated. As a measure for the performance and the complexity, we will employ the symbol error rate (SER) and the average number of nodes visited. The initial radius is $C = \infty$. For each SNR point, we run at least 10, 000 channel realizations.

Fig. 2 and Fig. 3 provide the performance and complexity results as a function of SNR for 8×8 MIMO system with 4-QAM when we choose ratio thresholds $\overline{\gamma}_1 = [\ 1\ 2\ 2\ 2\ 2\ 4\ 4\ 6\ 6\ 8\ 8\ 10\ 10\ \infty\ \infty\ \infty\ ]^T$, $\overline{\gamma}_2 = [\ 1\ 1.2\ 1.2\ 1.3$

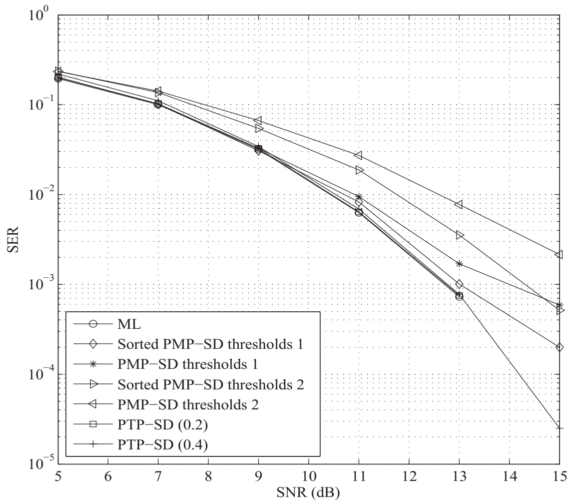

Fig. 4. Performance of PMP-SD, sort PMP-SD and PTP-SD for 8×8 MIMO system with 4-QAM

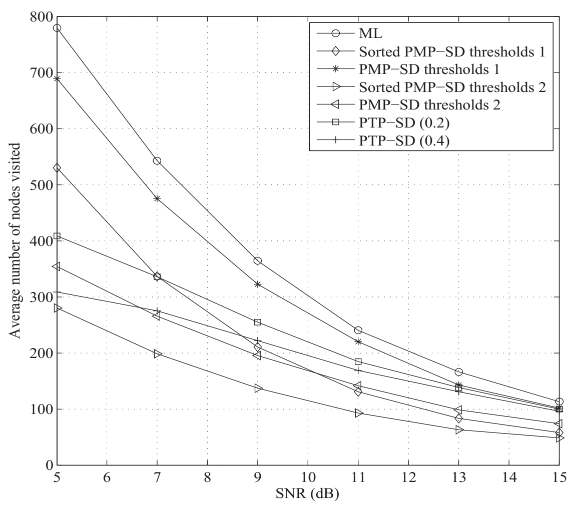

Fig. 5. Complexity of PMP-SD, sort PMP-SD and PTP-SD for 8×8 MIMO system with 4-QAM

1.3 1.4 1.4 1.5 1.5 2 4 8 12 ∞ ∞ $\infty]^T$, and $\overline{\gamma}_3 = [$ 1 1.2 1.2 1.3 1.3 1.4 1.4 1.5 1.5 2 4 8 12 16 20 24 $]^T$. It is obvious that the complexity is significantly reduced at the cost of different levels of performance loss. According to the line corresponding to $\overline{\gamma}_1$ and $\overline{\gamma}_2$, we can find that the tradeoff between performance and complexity can be easily achieved by adjusting the ratio thresholds. The larger the thresholds are, the less the loss of performance and reduction of complexity are, and vice versa. Lines corresponding to $\overline{\gamma}_2$ and $\overline{\gamma}_3$ present that it is important to choose different threshold for each layer. If we start pruning from the top layers, the reduction of complexity is negligible, but the degradation of performance is significant.

Fig. 4 and Fig. 5 present the influence of the sorted QR-decomposition. The thresholds $\overline{\gamma}_1$ and $\overline{\gamma}_2$ are used. Obviously, the sort can reduce the complexity and improve the performance simultaneously. Additionally, the two figures compare the PMP-SD with the PTP-SD [11]. The pruning probabilities defined in [11] are marked in the figures. We can find that PTP-SD has better performance but PMP-SD can reduce the complexity more significantly especially at the high SNR regime where PTP-SD have little effect on complexity reduction.

V. Conclusion

In this paper, we propose a pruning algorithm for SD based on path metric. The tradeoff between performance and complexity can be easily achieved by adjusting the ratio thresholds. Without consideration of the overhead of the preprocessing, the additional computational complexity which is just one comparison operation for one node can be ignored. The sorted QR-decomposition, which is optional, can simultaneously imporove the performance and the complexity, although more computation is needed at the stage of preprossing. Proposed methods can find wide applications in many areas such as multi-user detection.

References

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41-59, 1996.

[2] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky,"Detection algorithm and initial laboratory results using the V-BLAST space-time communication architecture," *Electronics Lett.*, vol. 35, pp.14-15, Jan. 1999.

[3] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics Computation*, vol. 44, pp. 463-471, 1985.

[4] C. P. Schnorr and M. Euchner, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Math. Programming*, vol. 66, pp. 181-191, 1994.

[5] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice codes decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, pp. 161-163, May 2000.

[6] M. O. Damen, H. E. Gamel, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2389-2402, Oct. 2003.

[7] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm-I: expected complexity," *IEEE Trans. Signal Proc.*, vol. 53, pp. 2806-2818, Aug. 2005.

[8] B. Shim and I. Kang, "On Further Reduction of Complexity in Tree Pruning Based Sphere Search," *IEEE Trans. Commun.*, vol. 58, pp. 417-422, Feb, 2010.

[9] A. Schenk, R. F. H. Fischer, and L. Lampe, "A Stopping Radius for the Sphere Decoder and its Application to MSDD of DPSK," *Communications Letters. IEEE*, vol. 13, no. 7, pp. 465-467, Jul. 2009.

[10] R. Gowaikar and B. Hassibi, "Statistical pruning for near-maximum likelihood decoding," *IEEE Trans. Signal Process.*, vol. 55, pp. 2661-2675, June 2007.

[11] B. Shim and I. Kang, "Sphere decoding with a probabilistic tree pruning," *IEEE Trans. Signal Process.*, vol. 56, pp. 4867-4878, Oct. 2008.

[12] D. Wübben, R. Böhnke, J. Rinas, V. Kühn and K.D. Kammeyer, "Efficient algorithm for decoding layered space-time codes," *IET Electronic Letters*, Vol. 37, No. 22, pp. 1348-1350, Oct. 2001