A Statistical Pruning Strategy for Schnorr-Euchner Sphere Decoding

Alireza Ghaderipoor, *Student Member, IEEE*, and Chintha Tellambura, *Senior Member, IEEE*

***Abstract*— The high computational complexity of maximum likelihood (ML) decoding can impact many applications such as code division multiple access (CDMA) and multiple-input multiple-output (MIMO) systems. The *sphere decoder* (SD) as an efficient ML decoder has therefore received significant attention in the wireless research community. This letter presents a new statistical method to reduce the complexity of the Schnorr and Euchner sphere decoder (SESD) [1]. The method uses a set of bounds, which are computed using the conditional probability based on the minimum metric of the current solution. A lookup table for the bounds can be computed offline. The proposed method is effective for any number of antennas with complexity savings about 50% or more over the conventional SD approach.**

***Index Terms*— MIMO, sphere decoding, closest point, lattice decoding.**

I. Introduction

IN code division multiple access (CDMA) and multiple antenna systems, optimum maximum likelihood (ML) decoding is, in general, NP-hard [1]. Nevertheless, in many cases, for a certain range of system parameters such as signal-to-noise ratio (SNR) and system dimension, the average complexity of some algorithms implementing or approximating ML decoding is polynomial in system dimensions. This fact has been shown by theoretical analysis [2] for the basic sphere decoder (SD) known as the Fincke-Pohst sphere decoder (FPSD) [3].

Recently, the basic SD has been modified to reduce the computational complexity, specially in low SNR, while keeping the system performance close to ML. Since the SD involves a tree search, the complexity reduction requires *pruning* the search tree. In [4], two statistical pruning methods have been proposed for the FPSD that are effective for *large dimensions* only. In both methods, instead of a single bound, a set of bounds is used for metric comparisons at all levels of the search tree. The more effective method of these two is called the *Increasing Radii Algorithm* (IRA).

This letter focuses on pruning methods for the Schnorr and Euchner SD (SESD). Since the SESD is more efficient than the basic FPSD [1], it makes sense to develop pruning strategies for the SESD. The SESD adopts an ordered set for the search of nodes of each level on the search tree, and this choice of the search order leads to computational efficiency improvements over the basic FPSD [1]. In the SESD, the cost metric of the current best solution (minimum metric) is used as a bound during the decoding. To reduce the complexity of SESD, we propose to limit the node metrics and/or differential metrics of two successive nodes of the search tree by using conditional probabilities based on the minimum metric. The simulation results show that our new statistically pruned (SP)-SESD reaches the ML performance in any arbitrary SNR region with a complexity less than other methods including IRA.

Manuscript received September 12, 2007. The associate editor coordinating the review of this letter and approving it for publication was J. Louveaux. This work has been supported in part by the Informatics Circle of Research Excellence (*i*CORE) and the Alberta Ingenuity Fund (AIF).

The authors are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada T6G 2V4 (e-mail: {ghaderi, chintha}@ece.ualberta.ca).

Digital Object Identifier 10.1109/LCOMM.2008.071518.

II. System Model and SD Algorithms

Consider the equivalent real multiple-input multiple-output (MIMO) linear model [5]:

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{z} \quad (1)$$

where $\mathbf{r} \in \mathbb{R}^M$ is the equivalent received vector, $\mathbf{z} \in \mathbb{R}^M$ is the equivalent additive noise vector, $\mathbf{H} \in \mathbb{R}^{M \times N}$ is the equivalent channel matrix and $\mathbf{s} \in \mathcal{Q}^N$ is the equivalent transmitted vector. The noise components are independent and identically distributed (i.i.d.) zero-mean Gaussian random variables with common variance σ^2, denoted by $\mathcal{N}(0, \sigma^2)$, the channel gains are i.i.d. $\mathcal{N}(0, 1/2)$, and the transmitted signals are selected from the set $\mathcal{Q} = \{2q - (Q-1) | q = 0, 1, ..., Q-1\}$. The SNR per transmitted symbol is defined as $\text{SNR} = E_s/\sigma^2$ where E_s is the average energy of the transmitted symbols.

We consider the QR decomposition of the channel matrix, $\mathbf{H} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q} \in \mathbb{R}^{M \times N}$ is an orthogonal matrix and $\mathbf{R} \in \mathbb{R}^{N \times N}$ is an upper triangular matrix with non-negative diagonal entries. The following equivalent MIMO model is then obtained:

$$\mathbf{y} = \mathbf{R}\mathbf{s} + \mathbf{v} \quad (2)$$

where $\mathbf{y} = \mathbf{Q}^{\mathrm{T}}\mathbf{r}$, and $\mathbf{v} = \mathbf{Q}^{\mathrm{T}}\mathbf{z}$ is an i.i.d. noise vector with the same statistical properties as $\mathbf{z}$. The ML decoder may be written as

$$\mathbf{x}_{\mathrm{ML}} = \arg \min_{\mathbf{x} \in \mathcal{Q}^N} \|\mathbf{y} - \mathbf{R}\mathbf{x}\|^2. \quad (3)$$

(We call $\|\mathbf{y} - \mathbf{R}\mathbf{x}_{\mathrm{ML}}\|^2$, the ML metric.)

The SD searches only over lattice points that lie in a hyper-sphere of radius $\sqrt{B}$ to reduce the search space. Thus, it examines all $\mathbf{x} \in \mathcal{Q}^N$ that satisfy

$$\|\mathbf{y} - \mathbf{R}\mathbf{x}\|^2 \leqslant B. \quad (4)$$

The cost metric in (3) may be expanded as

$$\begin{aligned} \|\mathbf{y} - \mathbf{R}\mathbf{x}\|^2 &= (y_N - r_{N,N}x_N)^2 \\ &+ (y_{N-1} - r_{N-1,N-1}x_{N-1} - r_{N-1,N}x_N)^2 \\ &+ \cdots + c(N) \leqslant B\,. \end{aligned} \quad (5)$$

Since the first term in (5) depends on x_N only, the second term depends on $\{x_N, x_{N-1}\}$ and so on, a tree can be constructed

1089-7798/08$25.00 © 2008 IEEE

with depth N. In the nth level of the tree, each node has a partial metric (cost) for $\{x_N, ..., x_{N-n+1}\}$ given by

$$c(n) = \left(y_{N-n+1} - \sum_{i=N-n+1}^{N} r_{N-n+1,i} x_i \right)^2 , \, n = 1, ..., N \tag{6}$$

and therefore the total metric of each nth level node is

$$C(n) = \sum_{i=1}^{n} c(i) \, , \, n = 1, ..., N. \tag{7}$$

For the metric of the actual transmitted vector $\mathbf{s}$, we use c_n and C_n instead of $c(n)$ and $C(n)$ respectively. The total metric of the actual transmitted vector is

$$\begin{aligned} C_N = \|\mathbf{y} - \mathbf{R}\mathbf{s}\|^2 &= v_N^2 + v_{N-1}^2 + v_{N-2}^2 + \cdots + v_1^2 \\ &= c_1 \; + c_2 \quad + c_3 \quad + \cdots + c_N \end{aligned} \tag{8}$$

where v_n's are the elements of $\mathbf{v}$ in (2).

The search tree has Q^N end nodes corresponding to all possible transmitted vectors. The SD discards unnecessary nodes by comparing the successive node metrics with the bound B. Two basic ideas have been proposed for the selection of B. The first one, the Fincke-Pohst (FP) strategy, is based on the fact that C_N (8) is a central chi-square random variable with N degrees of freedom, denoted by $\chi(N, \sigma^2)$. Thus, B is selected such that with a high probability, C_N is less than B. With this choice of B, if the algorithm fails to find at least one end node, B is increased and the algorithm is executed again. The process is repeated until at least one valid solution is found. The second strategy for selection of B, the Schnorr-Euchner (SE) strategy, uses the *QR (Babai) point* metric as the initial bound B [1]. The QR point is found by minimizing the metric terms in (5) successively. Whenever the SESD finds a new solution with a smaller metric, it is stored as the minimum cost solution $\mathbf{x}_{\min}$, and the new metric is substituted in $B_{\min}$. The SESD also uses a zigzag search order for x_n in (5) which makes it faster than FPSD (see [1] for more details).

III. Statistical Pruning

A. Pruning based on node metrics

Statistical pruning uses a set of bounds $\mathcal{B} = \{B_1, ..., B_n, ..., B_N\}$ instead of only one bound B, for node metrics examinations. The nth level node metrics, $C(n)$'s, are compared with B_n. Since C_n is a central chi-square random variable with n degrees of freedom, $\chi(n, \sigma^2)$, one way to choose $\mathcal{B}$ is based on the cdf of $\chi(n, \sigma^2)$, or equivalently B_n is chosen so that

$$\Pr\left[\, C_n \leqslant B_n \,\right] \geqslant P$$

where P is a predefined probability. Usually P is defined as:

$$P = P_\alpha \triangleq 1 - 10^{-\alpha} \quad , \quad \alpha = 1, 2, ... \tag{9}$$

This idea is known as the *increasing radii algorithm* (IRA) [4]. In the IRA with the FPSD, for a small α, $\mathcal{B}$ is computed. If the decoder fails to find an end node, α is increased and the algorithm is executed again.

Since the FPSD is less efficient than the SESD [5], we propose to use the set of bounds $\mathcal{B}$ with the SESD, expecting a decoder with less complexity than IRA with FPSD. The current minimum cost solution is $\mathbf{x}_{\min}$, and the corresponding metric is $B_{\min} = \|\mathbf{y} - \mathbf{R}\mathbf{x}_{\min}\|^2$. Since the ML metric is the minimum achievable metric during the decoding, i.e. $C(N) \leqslant B_{\min}$, the following conditional probability is used to determine $\mathcal{B}$:

$$\Pr\left[\, C_n \leqslant B_n \,|\, C_N \leqslant B_{\min} \,\right] \geqslant P_\alpha. \tag{10}$$

Note that (10) is reduced to IRA when $B_{\min} = \infty$.

Since C_n is $\chi(n, \sigma^2)$, the conditional probability can be evaluated as

$$\begin{aligned} &\Pr\left[\, C_n \leqslant B_n \,|\, C_N \leqslant B_{\min} \,\right] \\ &= \frac{\Pr\left[\, C_n \leqslant B_n \,,\, C_N \leqslant B_{\min} \,\right]}{\Pr\left[\, C_N \leqslant B_{\min} \,\right]} \\ &= \frac{\int_0^{B_n} \Pr\left[\, C_N \leqslant B_{\min} \,|\, C_n = y \,\right] f_{C_n}(y)\, dy}{\Pr\left[\, C_N \leqslant B_{\min} \,\right]} \\ &= \frac{\int_0^{B_n} \Pr\left[\, C_{N-n} \leqslant B_{\min} - y \,\right] f_{C_n}(y)\, dy}{\Pr\left[\, C_N \leqslant B_{\min} \,\right]} \\ &= \frac{\int_0^{B_n} F_{C_{N-n}}(B_{\min} - y) f_{C_n}(y)\, dy}{F_{C_N}(B_{\min})} \end{aligned} \tag{11}$$

where $f_{C_n}(y)$ and $F_{C_n}(y)$ are the pdf and cdf of $\chi(n, \sigma^2)$, respectively. Finally, from (10) and (11) and by a normalization, we obtain the following equation that can be used to find B_n as a function of $B_{\min}$:

$$\begin{aligned} P_\alpha = &\frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{N-n}{2})} \\ &\times \frac{\int_0^{a_1} \int_0^{a_2} u^{\frac{N-n}{2}-1} y^{\frac{n}{2}-1} \exp(-u-y) du dy}{\int_0^{a_3} y^{\frac{N}{2}-1} \exp(-y) dy} \end{aligned} \tag{12}$$

where $a_1 = \frac{B_n}{2\sigma^2}$, $a_2 = \frac{B_{\min}}{2\sigma^2} - \frac{y}{2\sigma^2}$, $a_3 = \frac{B_{\min}}{2\sigma^2}$ and $\Gamma(\cdot)$ is the Gamma function. To find the bound set $\mathcal{B}$, one can numerically solve (12), by using the normalized minimum metric $\frac{B_{\min}}{2\sigma^2}$.

B. Pruning based on differential metric

Another way to reduce the search over the tree is to limit the node metric increment from one node to its chid node. This requires a bound, B_d, for the differential metric, $c(n)$, defined in (6). Similar to (10), when the minimum metric $B_{\min}$ is determined, the following conditional probability can be used to find B_d:

$$\Pr\left[\, c_n \leqslant B_d \,|\, C_N \leqslant B_{\min} \,\right] \geqslant P_\alpha \tag{13}$$

Since c_n has $\chi(1, \sigma^2)$ distribution, the differential bound B_d can be extracted using equation (12) for $n = 1$ and normalized minimum metric $\frac{B_{\min}}{2\sigma^2}$. Since the probability is conditioned on the minimum metric, B_d must be updated whenever the minimum metric is changed during the decoding.

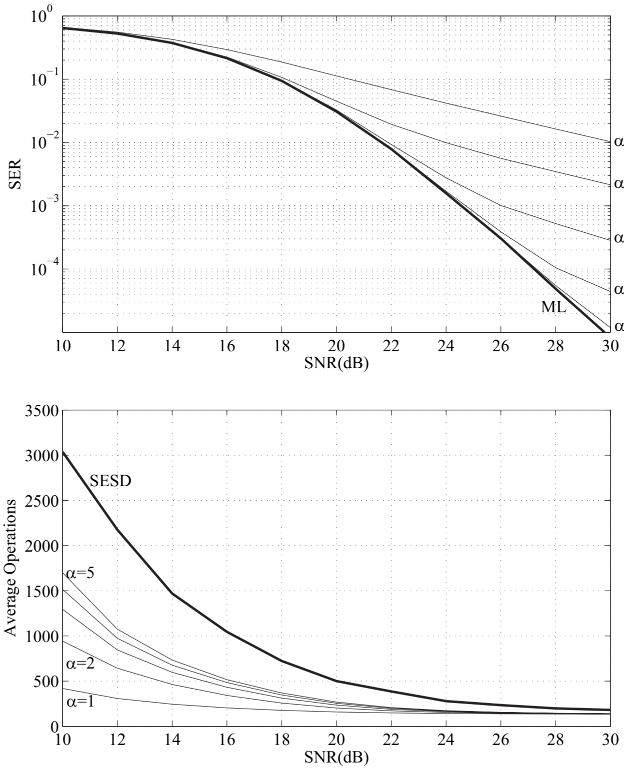

Fig. 1. Symbol error rate and average operations for MIMO(4,4,64) system using SP-SESD with $\alpha = 1, ..., 5$.

IV. Simulation Results

Since (12) depends on the choice of α and the normalized minimum metric $\frac{B_{\min}}{2\sigma^2}$, $\mathcal{B}$ can be pre-computed for some discrete values of the normalized minimum metric and α. For our simulations, we calculated $\mathcal{B}$ for $\alpha = \{1, 2, ..., 5\}$ and $\frac{B_{\min}}{2\sigma^2} = \{1, 1.5, 2, 2.5, ..., B_{\max}\}$ where $B_{\max}$ is calculated from $\Pr[\chi(N, 1/2) \leqslant B_{\max}] \geqslant P_{10}$. During the decoding, the minimum metric $B_{\min}$ is normalized to $\frac{B_{\min}}{2\sigma^2}$, then the set of bounds corresponding to the $\lceil \frac{B_{\min}}{\sigma^2} \rceil / 2$ is loaded and this set multiplied by $2\sigma^2$ is used as $\mathcal{B}$ where always $B_N = B_{\min}$ and $B_d = B_1$. For all simulations, an uncoded MIMO system with 4 transmit and 4 receive antennas is considered, where the transmitted symbols are selected from a 8^2-QAM constellation (we refer to this system as MIMO(4,4,64)). Considering the equivalent real system model (1), the channel matrix is $\mathbf{H} \in \mathbb{R}^{M \times N}$ where $M = N = 8$ and $\mathbf{x} \in \{\pm 7, \pm 5, \pm 3, \pm 1\}^8$.

Our simulation results show that pruning with the total metric (7) or the differential metric (6) have almost the same performance and complexity for different α. Consequently, statistical pruning is used with the differential metric and the SESD, because although both methods have the same complexity, the use of absolute node metrics needs more resources (memory) for implementation. Fig. 1 shows the performance and complexity of MIMO(4,4,64) system, using statistical pruning with differential metric, for different α. Since in this case $\lceil B_{\max} \rceil = 64$, we used a pre-calculated vector of 128 bounds for B_d. The simulation results show that for each α, the SP-SESD algorithm has the ML performance up to a certain SNR. When α is increased, the complexity of the SP-SESD is increased in low SNR, but still the complexities are close to each other for different values of α. The interesting point is that for a wide range of SNR, particularly for SNR$\geqslant$18 dB where the performance is acceptable for a wireless link, the complexity of the SP-SESD is almost independent of α.

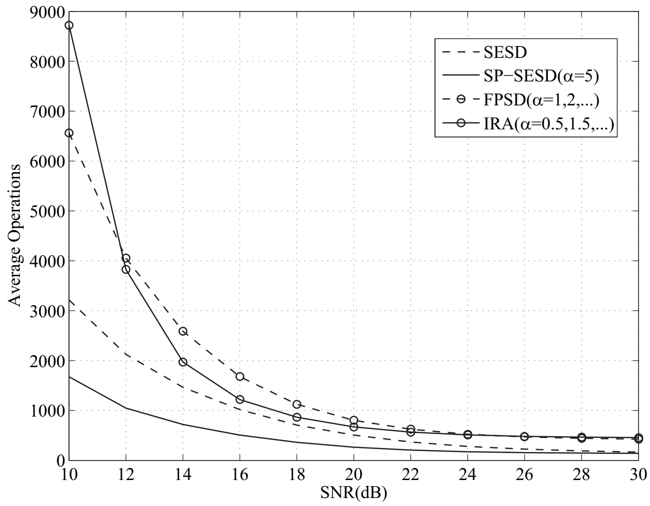

Fig. 2. Average operations for MIMO(4,4,64) system with different decoding methods.

Fig. 2 shows the average complexity of SP-SESD, SESD, FPSD and IRA algorithms for the decoding of a MIMO(4,4,64) system. As expected, in the whole SNR region, the SP-SESD has the minimum complexity compared to the others. The simulation results show that the average complexity of SP-SESD and SESD algorithms are almost 60% lower than FPSD and IRA at high SNR (SNR$\geqslant$24 dB). Fig. 2 shows that the complexity of the SESD has been reduced almost 50% in low to moderate SNRs by augmenting with statistical pruning.

V. Conclusion

This letter proposed a statistical pruning method for the SESD. The method uses a set of bounds which are computed using the conditional probability based on the minimum metric of the current solution. A lookup table for the bounds can be computed offline. The proposed method is effective for any number of antennas with complexity savings about 50% or more over the IRA (which is based on the FPSD), which tends to be effective for large dimensions only [4].

References

[1] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2201–2214, Aug. 2002.

[2] B. Hassibi and H. Vikalo, "On the expected complexity of sphere decoding," in *Proc. Asilomar Conf. on Signals, Systems and computers*, vol. 2, pp. 1051–1055, Nov. 2001.

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

[4] R. Gowaikar and B. Hassibi, "Efficient near-ml decoding via statistical pruning," in *Proc. IEEE International Symposium on Information Theory 2003*, pp. 274–274.

[5] M. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389–2402, 2003.