MIMO Cube Decoder

Masoud Alghoniemy and Ahmed H. Tewfik

Abstract—**An insight on the lattice decoder for flat-fading multiple antenna wireless communications systems is presented in this paper. In particular, we show that by formulating the decoding problem as a bounded-error subset selection, the resultant decoder finds the nearest lattice point to the received signal vector such that the search is bounded inside a hypercube centered at the received vector. The dimensions and orientation of the hypercube can be adjusted based on the diversity of the channel in order to improve its performance. The search for the nearest codeword to the received signal vector is solved by modeling the problem as an Integer Program (IP). Simulation shows that the proposed decoder is inferior to the Sphere Decoder (SD) by about 1-dB while its complexity is superior to the Sphere Decoder at very low signal to noise ratio.**

I. INTRODUCTION

Searching for the nearest lattice point to a given point in multidimensional lattices arises in Multi-Input Multi-Output (MIMO) wireless communication systems, e.g., Spatial Multiplexing (SM) and Space-Time (ST) Codes. When the contaminating noise is Gaussian, the optimum receiver, the Maximum Likelihood (ML), is the minimum distance receiver which leads to exhaustive search across all lattice points. Unfortunately, with MIMO systems the lattice size grows exponentially with the number of transmit antennas and the ML solution would be infeasible. In particular, consider the complex-valued baseband MIMO model in flat fading channels with M transmit and N received antennas. Let the $N \times 1$ received signal vector $\grave{y}$,

$$\grave{\mathbf{y}} = \grave{\mathbf{H}}\grave{\mathbf{x}} + \grave{\mathbf{w}} \tag{1}$$

with the transmitted signal vector $\grave{\mathbf{x}} \in \mathbb{Z}_c^M$ whose elements are drawn from q-QAM constellation and $\mathbb{Z}_c$ is the set of complex integers, the $N \times M$ channel matrix $\grave{\mathbf{H}}$ whose elements h_{ij} represent the Rayleigh complex flat fading gain from transmitter j to receiver i with $h_{ij} \sim \mathcal{CN}(0,1)$. In this paper, it is assumed that channel realization is known to the receiver through preamble and/or pilot signals, and $N \geq M$. The $N \times 1$ complex noise vector $\grave{\mathbf{w}}$ has independent complex Gaussian elements with variance σ^2 per dimension. Throughout the paper, we will consider the real model of (1)

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

M. Alghoniemy is with the Department of Electronics and Communications Engineering, Egypt-Japan University of Science and Technology, Egypt (e-mail: masoud.alghoniemy@ejust.edu.eg). He is also with the Department of Electrical Engineering, University of Alexandria, 21544 Egypt.

A. H. Tewfik is with the Department of Electrical and Computer Engineering, University of Texas at Austin, TX, USA (e-mail: tewfik@austin.utexas.edu)

Define $m = 2M, n = 2N$, then $\mathbf{y} = [\mathcal{R}(\grave{\mathbf{y}}) \ \mathcal{I}(\grave{\mathbf{y}})]^T \in \mathbb{R}^n$, $\mathbf{x} = [\mathcal{R}(\grave{\mathbf{x}}) \ \mathcal{I}(\grave{\mathbf{x}})]^T \in \mathbb{Z}^m$, $\mathbf{w} = [\mathcal{R}(\grave{\mathbf{w}}) \ \mathcal{I}(\grave{\mathbf{w}})]^T \in \mathbb{R}^n$, and

$$\mathbf{H} = \left(\begin{array}{cc} \mathcal{R}(\grave{\mathbf{H}}) & -\mathcal{I}(\grave{\mathbf{H}}) \\ \mathcal{I}(\grave{\mathbf{H}}) & \mathcal{R}(\grave{\mathbf{H}}) \end{array} \right) \in \mathbb{R}^{n \times m},$$

where $\mathcal{R}(.)$ and $\mathcal{I}(.)$ are the real and imaginary parts, respectively. ML solution finds the symbol estimate $\hat{\mathbf{x}}_{\mathbf{ML}}$ that minimizes the 2-norm of the residual error. In particular

$$\hat{\mathbf{x}}_{\mathbf{ML}} = arg \min_{\mathbf{x} \subset \Lambda} \parallel \mathbf{y} - \mathbf{H}\mathbf{x} \parallel_2 \tag{3}$$

where Λ is the lattice whose points represent all possible codewords at the transmitter and $\mathbb{Z}^m$ is the set of integers of dimension m. The coordinates of the lattice points are all integers whose elements are drawn from the $L = \log_2(q)$-PAM derived from the q-QAM constellation. According to (2), ML solution leads to solving integer least-squares problem which is, in general, NP-hard [1]. Moreover, the number of lattice points in a given lattice Λ can be extremely large even for a reasonable number of transmit antennas. In SM for example, if $M = 4$ and 16-QAM constellation is used, then the number of all possible lattice points is 16^4 codewords. Another difficulty comes from the fact that the elements of $\hat{\mathbf{x}}$ are restricted to be integers and can only take values from the corresponding L-PAM.

Approximate solutions to (3) include the Zero Forcing (ZF) and MMSE solutions, V-BLAST which is a Nulling and Canceling technique with or without optimal ordering [2], [3]. Exact methods with reduced complexity is found for special orthogonal Space-Time Block Codes (STBC) [4], [5]. The Sphere Decoder (SD) of Fincke and Pohst searches for the lattice point inside a sphere centered at the received vector [6]–[8]. The SD achieves near-optimum performance but its complexity is a function of the SNR; a major drawback of the sphere decoder is that its average complexity is high for low signal to noise ratio and decreases with high SNR; this makes it difficult to use for low SNR environments [9].

The contribution of this paper can be summarized in the following two points. First, we show that by reformulating the decoding problem as a bounded-error subset selection, the resultant decoder finds the nearest lattice point to the received vector inside a hypercube centered at that point, hence the name "Cube Decoder" (CD). We will show that the dimensions of the hypercube depends on channel diversity. In particular, for diverse channel, the channel matrix is well-conditioned and the cube will not be skewed much. On the other hand, for non-diverse channel, the channel matrix is ill-conditioned and the cube will be highly skewed. The second contribution, we show by simulation that the complexity of the proposed

decoder is lower than the SD for very low SNR which makes it suitable for low SNR environments.

II. THE BOUNDED-ERROR SUBSET SELECTION

We will use the bounded-error subset selection problem for deriving the cube decoder. The bounded-error subset selection was introduced by the authors in [10], [11] which we review here for completeness. In the subset selection problem it is required to find the best, according to some criterion, signal representation for a signal vector $\mathbf{y}$ using an overcomplete dictionary represented by the N dimensional vectors spanning the column space of the matrix $\mathbf{H}$. By construction, the number of basis vectors M in the dictionary is such that $M \gg N$. In case of sparseness criterion, it is required to find the sparsest vector $\mathbf{x}$ (the vector $\mathbf{x}$ with the minimum number of non-zero coefficients) such that

$$\mathbf{y} \simeq \mathbf{Hx}. \tag{4}$$

It is known that the subset selection problem is NP-hard [1]; there are several approximate solutions to the subset selection problem with low complexity (e.g., [12]–[15]). The bounded-error subset selection finds the sparsest solution vector $\mathbf{x}$ by relaxing the equality constraint and introducing a bounded version of (4)

$$\mathbf{y}_{min} \leq \mathbf{Hx} \leq \mathbf{y}_{max} \tag{5}$$

where, $\mathbf{y}_{min} = \mathbf{y} - \vec{\epsilon_1}$ and $\mathbf{y}_{max} = \mathbf{y} + \vec{\epsilon_2}$. Here, $\vec{\epsilon_1}$ and $\vec{\epsilon_2}$ are error vectors that could represent model uncertainties. The relaxation introduced in the bounded-error subset selection allows more degrees of freedom which can be used to find the sparse solution to (5). Sparseness is imposed by minimizing the number of non-zero elements in the solution vector $\mathbf{x}$. This can be achieved as follows: if $x_k \in \{0, 1\}$, then the number of non-zero elements in $\mathbf{x}$ is $\sum_k x_k = \mathbf{1}^T \mathbf{x}$ where $\mathbf{1}$ is a vector of all ones. Hence finding a sparse solution leads to following binary integer program

$$\begin{aligned} \min_{x_k \in \{0,1\}} \quad & \mathbf{1}^T \mathbf{x} \\ s.t. \quad & \mathbf{y}_{min} \leq \mathbf{Hx} \leq \mathbf{y}_{max} \end{aligned} \tag{6}$$

The integer program (6) is called the bounded-error subset selection.

III. THE CUBE DECODER

In this section, we develop the cube decoder by formulating the search problem as a bounded-error subset selection. It should be noted that the objective function, $\mathbf{1}^T \mathbf{x}$, in (6) is equivalent to $\parallel \mathbf{x} \parallel_1$ because x_k assume only non-negative values. Hence, the bounded error subset selection (6) can be re-written as

$$\begin{aligned} \min_{x_k \in \{0,1\}} \quad & \parallel \mathbf{x} \parallel_1 \\ s.t. \quad & \mathbf{y}_{min} \leq \mathbf{Hx} \leq \mathbf{y}_{max} \end{aligned} \tag{7}$$

Since in decoding problems we are only interested in finding the transmitted codeword $\mathbf{x}$ that is nearest, in some sense, to the received signal $\mathbf{y}$; then solving a feasibility optimization problem would suffice to get an estimate to the transmitted codeword. Feasibility optimization problems finds the solution

that fulfills the constraints regardless of the objective function. For example, the constraint in (7) guarantees that the estimated codeword $\mathbf{x}$ will be bounded, in some sense as we will see later, to the received signal $\mathbf{y}$. However, solving only a feasibility problem would not reduce the estimated error because this has not been considered in the optimization. In order to improve the estimation performance, minimizing an objective function of the error would reduce the estimation error. In this paper, we consider the ℓ_1 norm of the residual $\mathbf{r} = \mathbf{y} - \mathbf{Hx}$ as an objective function; the choice of this particular objective function was made based on the following: (1) the corresponding decoding problem reduces to bounded-error subset selection with known solution. (2) it turns out that modeling the decoding problem as a bounded-error subset selection has an interesting geometric interpretation, that is the search space lies inside a hypercube centered at the received signal, hence the name "cube decoder". This has the advantage of understanding the behavior of the decoder and provides space for improvement.

That is, the proposed decoder finds the nearest codeword, in the ℓ_1 sense, to the received signal vector. In order to limit the search space, the received signal vector $\mathbf{y}$ is lower and upper bounded by $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$, respectively. The estimated codeword $\hat{\mathbf{x}}_{\ell_1}$ can be found by solving

$$\begin{aligned} \hat{\mathbf{x}}_{\ell_1} = arg \min_{\mathbf{x} \subset \Lambda} \quad & \parallel \mathbf{y} - \mathbf{Hx} \parallel_1 \\ s.t. \quad & \mathbf{y}_{min} \leq \mathbf{Hx} \leq \mathbf{y}_{max} \\ & \mathbf{x} \in \mathbb{Z}^m. \end{aligned} \tag{8}$$

where Λ is the lattice whose points represent all possible codewords at the transmitter, $\mathbb{Z}^m$ is the set of integers of dimension m. $\mathbf{y}_{min} = \mathbf{y} - \vec{\epsilon_1}$ and $\mathbf{y}_{max} = \mathbf{y} + \vec{\epsilon_2}$, with $\vec{\epsilon_1}$ and $\vec{\epsilon_2}$ are error vectors that depend on the channel matrix $\mathbf{H}$. The generalized vector inequality is element-wise inequality. It should be noted that the proposed decoder (8) follows the bounded-error subset selection model (7). However ℓ_1 based decoding is not new and has been explored before, e.g., [16], [17]. Here we will provide more insight for such decoding criterion.

A. Geometric Interpretation

The bounded-error subset selection formulation has an interesting geometric interpretation which is introduced in this section. We will show that the constraint in (8) represents a bound on the search space and this bound is in fact a hypercube. In order to show that, we can see that $\mathbf{Hx}$ represents a skewed lattice, and it is required to find the nearest lattice point, in the ℓ_1-norm sense, to the received signal vector $\mathbf{y}$. On the other hand, $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$ represent perturbed versions of $\mathbf{y}$.

By imposing the constraint $\mathbf{y}_{min} \leq \mathbf{Hx} \leq \mathbf{y}_{max}$ into (8), we limit the search space such that the estimated codeword $\hat{\mathbf{x}}_{\ell_1}$ is allowed to take only values such that $\hat{\mathbf{y}} = \mathbf{H}\hat{\mathbf{x}}_{\ell_1}$ lies between $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$. Figure 1 illustrates the points of the skewed lattice in 2-dimensional space, the received signal vector $\mathbf{y}$, the lower and upper bounds $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$, respectively. It is clear from the figure that $\mathbf{y}$ is located inside the rectangle whose two vertices are $\mathbf{y}_{min}$ and $\mathbf{y}_{max}$. Any
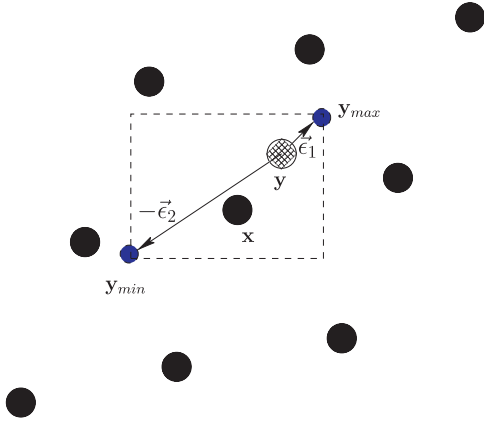
Fig. 1.　Geometric Interpretation in 2-D.

lattice point inside this rectangle would satisfy the constraints in (8). Hence, only lattice points inside the rectangle can be declared as possible solutions. In general, since the lattice under consideration is symmetric along each dimension, then one can always assume that $\vec{\epsilon}_1 = \vec{\epsilon}_2 = \vec{\epsilon}$. Hence, the rectangle reduces to a square in 2-D, a cube in 3-D and to a hypercube in multi-dimensions.

The choice of $\vec{\epsilon}$: is critical and could affect the performance of the decoder dramatically. This can be understood from figure 1 where $\vec{\epsilon}$ determine the size of the bounding cube and in higher dimensions it also determines its orientation as well. In particular, if the bounds in (8) are tight, which corresponds to small cube, then it is possible that the corresponding hypercube does not contain any lattice points inside and then there would be no solution. On the other hand, if the bounds are loose, then there would be too many lattice points inside the hypercube and it would take longer time to find the corresponding lattice point. Hence, the values of $\vec{\epsilon}$ determine the performance of the decoder. Since $\mathbf{H}$ can be viewed as a transformation matrix that transforms the original lattice into a skewed one, then it is natural to set $\vec{\epsilon}$ as the transformed lattice bases vectors. Hence,

$$\vec{\epsilon} = abs(\mathbf{H})(\vec{v_1} + \vec{v_2} + \cdots + \vec{v_m}) \tag{9}$$

where $\vec{v_k}$ is the kth lattice basis vector, and $abs(.)$ is element-wise absolute. Other choices of the error vectors $\vec{\epsilon}$ are possible and could be investigated in future research.

Since the lattice under consideration is finite, for low SNR it is legitimately possible that the received signal vector and the surrounding cube would lie completely outside the lattice. In this case, there would be no lattice points located inside the hypercube, and hence a failure is declared. In this case, the second constraint in (8) could be relaxed and unconstrained optimization is solved instead. This fact will be illustrated in the simulation section where we will show that for very low SNR, the complexity of the cube decoder is decreased due to solving unconstrained minimization problem which has lower complexity than its constrained counterpart.

B. Integer Program Solution

In this section, we will transform (8) into an integer program form [18], [19]. Let $\mathbf{r} = \mathbf{r}^+ - \mathbf{r}^-$ where $\mathbf{r}^+, \mathbf{r}^-$ are such that positive values of $\mathbf{r}$ goes to $\mathbf{r}^+$ while negative values goes to $\mathbf{r}^-$ but with positive sign. Similarly, let $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$. Define the $2(m+n) \times 1$ auxiliary vector $\mathbf{u} = [\mathbf{r}^+, \mathbf{r}^-, \mathbf{x}^+, \mathbf{x}^-]^{\mathbf{T}}$, the equality matrix $\mathbf{A}_{eq} = [\mathbf{I}, -\mathbf{I}, \mathbf{H}, -\mathbf{H}]$, and the non-equality matrix $\mathbf{A}_{neq} = [\mathbf{0}, \mathbf{0}, \mathbf{H}, -\mathbf{H}]$ where $\mathbf{I}$ and $\mathbf{0}$ are the identity and the zero matrix of size n. Then (8) can be re-written as

$$\begin{aligned} \hat{\mathbf{u}}_{\ell_1} = arg \min_{\mathbf{u}} \quad & \mathbf{f^T u} \\ s.t. \quad & \mathbf{y}_{min} \le \mathbf{A}_{neq}\mathbf{u} \le \mathbf{y}_{max} \\ & \mathbf{r}^+, \mathbf{r}^- \ge 0 \\ & \mathbf{x}^+, \mathbf{x}^- \in L^+ PAM \end{aligned} \tag{10}$$

with $\mathbf{f} = [\mathbf{1}_{2n}; \mathbf{0}_{2m}]$. $\mathbf{1}_{2n}$ and $\mathbf{0}_{2m}$ are vectors of ones and zeros of size $2n$ and $2m$, respectively. L^+-PAM represents only positive integers in the L-PAM set. For example, if q-QAM constellation is used and the corresponding L-PAM is represented as $\{-(L-1), -(L-3), \cdots, (L-3), (L-1)\}$ where $L = \log_2(q)$, then the elements of $\mathbf{x}^+$ and $\mathbf{x}^-$ can only take values from $L^+ = \{1, 3, .., L-1\}$, i.e., odd positive integers. A change of variables is necessary in order to remove the odd constraint. Let $x = 2z - (L-1)$, $\mathbf{z}^+$ and $\mathbf{z}^-$ are defined similarly, then (10) simplifies to

$$\begin{aligned} \hat{\tilde{\mathbf{u}}}_{\ell_1} = arg \min_{\tilde{\mathbf{u}}} \quad & \mathbf{f^T \tilde{u}} \\ s.t. \quad & \tilde{\mathbf{A}}_{eq}\tilde{\mathbf{u}} = \tilde{\mathbf{y}} \\ & \tilde{\mathbf{y}}_{min} \le \tilde{\mathbf{A}}_{neq}\tilde{\mathbf{u}} \le \tilde{\mathbf{y}}_{max} \\ & \mathbf{r}^+, \mathbf{r}^- \ge 0 \\ & \mathbf{z}^+, \mathbf{z}^- \in \mathbb{Z}^+ \end{aligned} \tag{11}$$

where $\tilde{\mathbf{u}} = [\mathbf{r}^+, \mathbf{r}^-, \mathbf{z}^+, \mathbf{z}^-]^{\mathbf{T}}$. $\tilde{\mathbf{A}}_{eq}$ and $\tilde{\mathbf{A}}_{neq}$ are such that $\mathbf{H}$ is replaced by $2\mathbf{H}$ in $\mathbf{A}_{eq}$ and $\mathbf{A}_{neq}$, respectively. Let the $m \times 1$ vector $\mathbf{1}_m$ be a vector of all ones, then

$$\tilde{\mathbf{y}} = \mathbf{y} + (L-1)\mathbf{H}\mathbf{1}_m \tag{12}$$

Or equivalently one can say that $\tilde{\mathbf{y}}_{max} = \tilde{\mathbf{y}} + \vec{\epsilon}$ and $\tilde{\mathbf{y}}_{min} = \tilde{\mathbf{y}} - \vec{\epsilon}$. Now, since $\mathbf{z}$ can only take positive integer values, then the lattice under consideration has bases vectors $\vec{v}_k = [0 \cdots 0 \ 1 \ 0 \ \cdots 0]^T$ where 1 is at the kth location. According to (9), $\vec{\epsilon} = 2 \, abs(\mathbf{H})\mathbf{1}_m$, where the factor 2 is due to the transformation of variables. It should be noted that in order to reduce the complexity of the IP, relaxation techniques could be considered [16], [19].

IV. SIMULATION

For comparison purposes, the performance of the proposed decoder is compared to the ML, SD, ZF and the MMSE decoders for the uncoded systems. It is assumed that the transmitted power is independent of the number of transmit antennas, M, and equals to the average symbols energy. The package, lp_solve, was used to solve (11) [20]. Figure 2 illustrates the performance of the proposed decoder, named CD decoder, for 4-QAM modulation for $N = M = 4$. As it is clear from the figure, the sphere decoder has the best performance which almost coincides with the ML decoder. However, the

Fig. 2. Average SER for 4×4.

(a) Average search time vs. SNR

(b) Average search time vs. QAM size
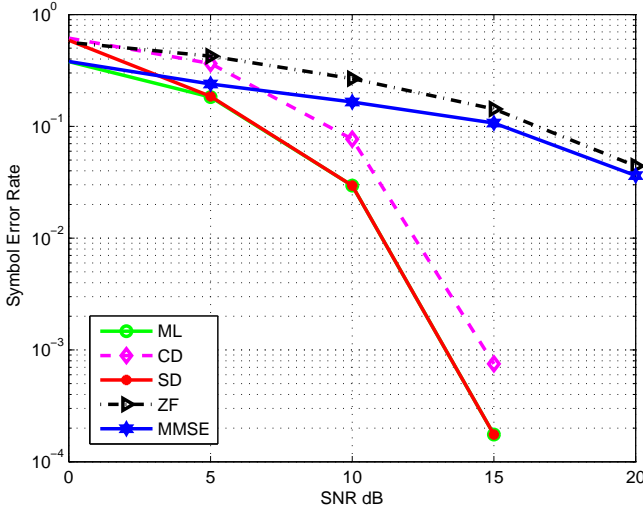
Fig. 3. Complexity for 4×4

performance of the proposed decoder is almost 1-dB inferior to the ML decoder and performs much better than both the ZF and MMSE decoders. This could be understood from the fact that in the Gaussian noise, ℓ_1-norm is not the optimal decoding criterion.

The complexity of the CD decoder is measured as the average time in seconds consumed by the solver in finding the nearest lattice point to the received vector. Figure 3 illustrates the complexity for 4×4 MIMO. In particular, figure 3(a) illustrates the average search time as a function of the SNR for different modulation schemes. It is clear that increasing SNR does not reduce the complexity of the decoder by a big margin. This is due to the fact that in IP, the complexity is much determined by the initialization of the algorithm, rather than SNR. This fact can be better understood from figure 3(b) which shows the average search time as a function of the constellation size for different SNRs where increasing the SNR from 10 to 20 dBs reduces the search time by almost 0.01 secs. However, on the contrary to the SD, the complexity of the CD decreases for very low signal to noise ratio as illustrated by figure 4. This reduction in complexity can be understood from the fact that for very low SNR, the cube decoder solves unconstrained optimization problem which lowers the search complexity.

V. CONCLUSION

In this paper we have examined the performance and the structure of the ℓ_1 lattice decoder in a bounded error subset selection formulation, for multi-antenna wireless systems. In particular, we have shown that by using the bounded-error formulation, the proposed decoder finds the nearest lattice-point inside a hypercube centered at the received signal vector using a Mixed Integer Linear Program. On the other hand, we also have shown that the proposed decoder is almost 1-dB inferior to ML-performance, while its complexity does not depend on the operating SNR; which makes it suitable for low SNR environments. Future areas of investigation could
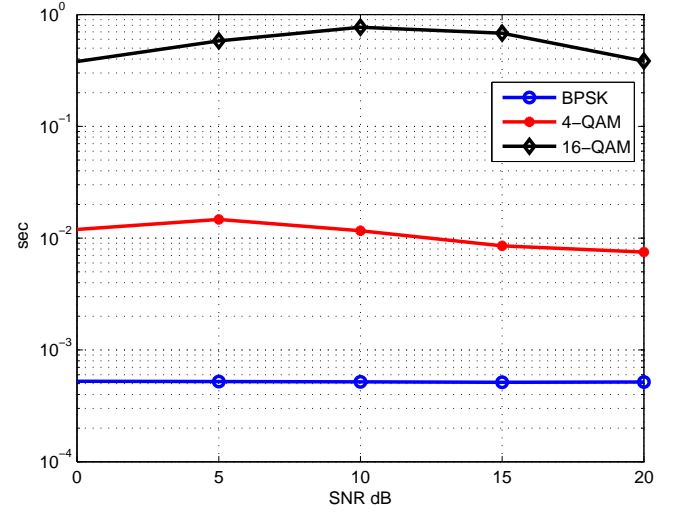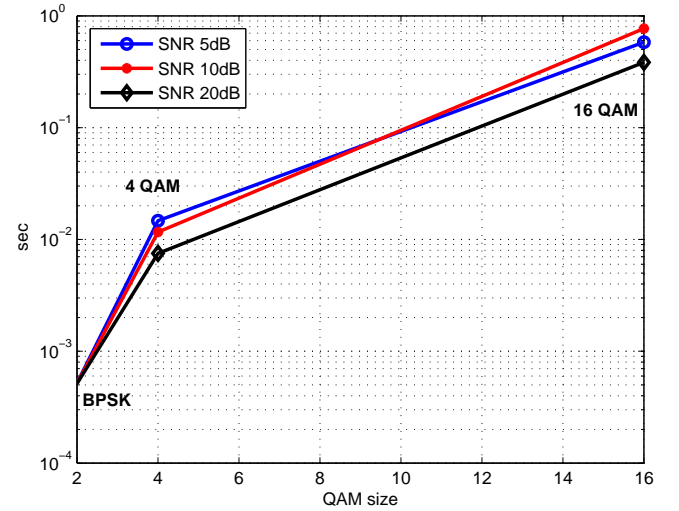
include: (1) an optimal choice of the cube dimensions and its orientation that better reflect channel conditions and (2) applying relaxation techniques for complexity reduction.

REFERENCES

[1] B. Natarajan, "Sparse Approximate Solutions to Linear Systems," *SIAM J. Comp.,* Vol. 24, pp. 227-234, Apr. 1995.
[2] G. J. Foschini, "Layered spacetime 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.
[3] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST spacetime communication architecture," *Electronics Letters,* vol. 35, pp. 1416, Jan. 1999.
[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Spacetime codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Information Theory,* vol. 44, pp. 744-765, Mar. 1998.
[5] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Spacetime block codes from orthogonal designs," *IEEE Trans. on Information Theory,* vol. 45, pp. 1456-1467, July 1999.
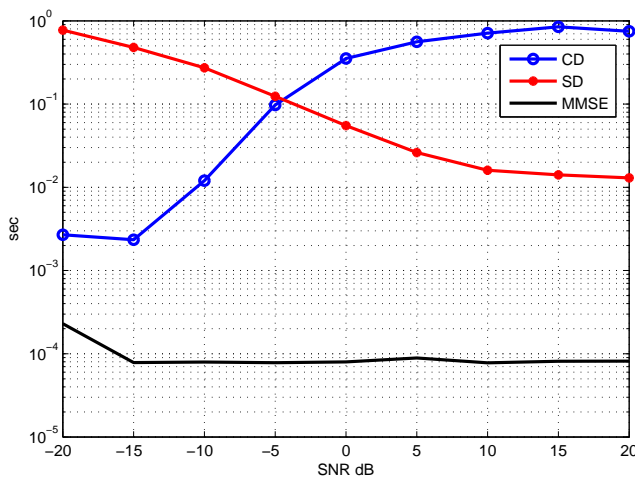
Fig. 4.   Complexity comparison for 4×4.

Masoud Alghoniemy received his B.Sc. and M.Sc. degrees from the University of Alexandria, Alexandria, Egypt, and the Ph.D. degree from the University of Minnesota, Minneapolis, in 1993, 1996, and 2001, respectively. He is an Associate Professor at the Egypt-Japan University of Science and Technology, Egypt. He was a Senior Audio Coding Engineer at iBiquity Digital Corporation, Warren, NJ, from April 2001 to March 2002. He was with Intel Corporation, Santa Clara, CA, as an intern in the summer of 2000. From 2002 to 2010, he has been an Assistant Professor with the Department of Electrical Engineering, University of Alexandria. He spent the period from October 2006 to October 2008 at the Malaysian Institute of Microelectronic Systems (MIMOS BHD), Malaysia, where he was part of a team responsible for implementing the PHY layer of the mobile WiMAX standard IEEE 802.16e. He worked as a senior research scientist at Intel Corp., Egypt from March - July 2010 where he worked on IEEE 802.16m.

His research interests include wireless communications and sparse signal representations. Dr. Alghoniemy was the recipient of the Young Scientist Award from the General Assembly of the International Union of Radio Science (URSI), Lille, France, in 1996.

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

[7]   Murugan, A. D.; Gamal, H. E.; Damen, M. O.; Caire, G. " A unified frame work for tree srach decoding: rediscovering the sequential decoder," *IEEE Trans. on Information Theory,* Vol. 52, No. 3, March 2006.

[8]   Agrell, E.; Eriksson, T.; Vardy, A.; Zeger, K. "Closest Point Search in Lattices," *IEEE Trans. on Information Theory,* Vol. 48, No. 8, Aug. 2002.

[9]   Haris Vikalo and Babak Hassibi, "On the Sphere-Decoding Algorithm II. Generalizations, Second-Order Statistics, and Applications to Communications," *IEEE Trans. on Signal Processing,* pp. 2819-2834, Vol. 53, No. 8, Aug. 2005.

[10] Masoud Alghoniemy and Ahmed H. Tewfik, "A Sparse Solution to the Bounded Subset Slection Problem: A Network Flow Model Approach," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing.,* Vol. 5, pp. 89-92, May 2004.

[11] M. Alghoniemy and Ahmed H. Tewfik, "Bounded Subset Selection with Noninteger Coefficients," *Proc. of the EUSIPCO Conferecne,* pp. 317-320, Sept. 2004.

[12] Chen S., and Donoho D., Saunders M. A., "Atomic Decomposition By Basis Pursuit," *SIAM J. on Scientific Computing,* Vol. 20, No. 1, pp. 33-61, 1998.

[13] Mallat S., and Zhang Z., "Matching Pursuit with Time-Frequency Dectionaries," *IEEE Trans. on Signal Processing,* Vol. 41, No. 12, pp. 3397-3415, Dec. 1993.

[14] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," *in Proc. 27th Asilomar Conference on Signals, Systems and Computers, A. Singh, ed.*, IEEE Comput. Soc. Press, Los Alamitos, CA, 1993.

[15] Coifman R., and Wickerhauser M., "Entropy-based Algorithms for Best Basis Selection," *IEEE Trans. on Info. Theory,* Vol. 38, No. 2, pp.713-718, March 1992.

[16] Cui, T.; Huo, T.; Tellambura, C. "Linear Programming Detection and Decoding for MIMO Systems," *Internationl Symposium on Information Theory (ISIT),* 2006.

[17] Candes, E.; Tao, T. "Decoding by Linear Programming," *IEEE Trans. on Information Theory,* Vol. 51, No. 12, Dec. 2005.

[18] S. Sardy, Regularization techniques for linear regression with a large set of carriers, Ph.D. dissertation, Univ. Washington, Seattle, 1998.

[19] Stephen Boyd and Lieven Vandenberghe, "Convex Optimization", *Cambridge University Press,* 2004.

[20] Michel Berkelaar, Kjell Eikland, Peter Notebaert, "Open source (Mixed-Integer) Linear Programming system: lpsolve" under GNU LGPL (Lesser General Public Licence).

Ahmed H Tewfik received his B.Sc. degree from Cairo University, Cairo Egypt, in 1982 and his M.Sc., E.E. and Sc.D. degrees from MIT, in 1984, 1985 and 1987 respectively. He is the Cockrell Family Regents Chair in Engineering and the Chairman of the Department of Electrical and Computer Engineering at the Univeristy of Texas Austin. He was the E. F. Johnson professor of Electronic Communications with the department of Electrical Engineering at the University of Minnesota until September 2010. Dr. Tewfik worked at Alphatech, Inc. and served as a consultant to several companies. From August 1997 to August 2001, he was the President and CEO of Cognicity, Inc., an entertainment marketing software tools publisher that he co-founded, on partial leave of absence from the University of Minnesota. His current research interests are in minimally invasive surgery, genomics and proteomics,, brain computing interfaces and programmable wireless networks. Prof. Tewfik is a Fellow of the IEEE. He was a Distinguished Lecturer of the IEEE Signal Processing Society in 1997 - 1999. He received the IEEE third Millennium award in 2000. He was elected to the position of VP Technical Directions of the IEEE Signal Processing Society in 2009 and served on the board of governors of that Society from 2006 to 2008. He has given several plenary and keynote lectures at IEEE conferences.