

Regularized MIMO Decoders

Masoud Alghoniemy

Abstract—In the Multi Input Multi Output (MIMO) antenna system, it is known that the Linear Minimum Mean Squared Error (MMSE) receiver is equivalent to Tikhonov regularization. Given that, we develop a family of generalized receivers based on regularization with different penalty functions that penalize the received symbols outside the convex hull of the modulating constellation. For illustration purposes we consider two types of penalty functions, the deadzone and infinity norm penalty functions. The proposed decoders have low complexity and can be implemented efficiently using convex optimization algorithms. Simulation results show that the proposed receivers outperform the MMSE receiver by as high as 5-dB at low Signal to Noise Ratio (SNR).

Index Terms—Multiple antenna systems, spatial multiplexing, lattice problems, wireless communications, regularization.

I. INTRODUCTION

In Spatial Multiplexing (SM) scenario of the Multi Input Multi Output (MIMO) flat fading wireless communication system with m transmit and n receive antennas, the relation between the transmitted and the received signal can be described as follows

$$\mathbf{y} = \sqrt{\frac{\rho}{m}} \mathbf{H} \mathbf{x} + \mathbf{w} \quad (1)$$

where ρ is the expected value of the Signal to Noise Ratio (SNR) at each receive antenna, \mathbf{x} is the $m \times 1$ transmitted vector whose elements are complex symbols drawn from the normalized M-QAM constellation with $\mathbf{E}(\mathbf{x}\mathbf{x}^T) = \mathbf{I}$, where M is the constellation order and \mathbf{I} is the $m \times m$ identity matrix. \mathbf{y} is the $n \times 1$ received vector, \mathbf{H} is the $n \times m$ channel matrix with $n \geq m$, whose elements represent the i.i.d. flat fading channel gains $h_{ij} \sim \mathcal{CN}(0, \sigma_h^2)$. Without loss of generality, \mathbf{w} is $n \times 1$, i.i.d. zero mean complex white Gaussian noise, uncorrelated with the transmitted symbols, with $w_i \sim \mathcal{CN}(0, 1)$. For the Gaussian noise scenario, the optimum decoder is the Maximum Likelihood decoder which finds the most likely input vector \mathbf{x}_{ml} according to

$$\mathbf{x}_{\text{ml}} = \arg \min_{\mathbf{x} \in \Lambda} \|\mathbf{y} - \sqrt{\rho/m} \mathbf{H} \mathbf{x}\|^2 \quad (2)$$

where Λ is the lattice whose points represent all possible combinations of \mathbf{x} . The problem (2) is NP-hard in general that can only be exactly solved by exhaustive search over all possible M^m vector combinations; where its complexity in this case grows exponentially with the problem size [1]. This is due to the discrete nature of the lattice Λ ; nearest lattice

point search algorithms can be used to solve (2) approximately [2], [3].

Hence, linear decoders have been used to obtain an approximate solution with lower complexity. The simplest linear decoder is the decorrelator which is known as the Zero-Forcing (ZF) decoder in which the constraint $\mathbf{x} \in \Lambda$ is relaxed and the domain in this case is \mathbb{R}^n . The zero-forcing decoder inverts the channel in order to cancel spatial interference, in particular $\mathbf{x}_{\text{zf}} = \mathbf{G}_{\text{zf}} \mathbf{y}$ with

$$\mathbf{G}_{\text{zf}} = \sqrt{m/\rho} (\mathbf{H}^* \mathbf{H})^{-1} \mathbf{H}^* \quad (3)$$

is the pseudo-inverse of the channel [4]. Although the zero forcing solution completely cancels out spatial interference, it has the disadvantage of enhancing the noise, especially if the channel matrix is ill-conditioned; in such case, small eigen values amplify the contaminating noise. In an effort to reduce noise enhancement, the Linear Minimum Mean Squared Error (MMSE) decoder is used to strike a balance between interference cancellation and noise enhancement [5]. The MMSE decoder finds the solution $\mathbf{x}_{\text{mmse}} = \mathbf{G}_{\text{mmse}} \mathbf{y}$ where

$$\mathbf{G}_{\text{mmse}} = \arg \min_{\mathbf{x}} \mathbf{E} \|\mathbf{G} \mathbf{y} - \mathbf{x}\|^2 \quad (4)$$

which has the analytical solution

$$\mathbf{G}_{\text{mmse}} = \sqrt{\frac{\rho}{m}} \left(\frac{\rho}{m} \mathbf{H}^* \mathbf{H} + \mathbf{I} \right)^{-1} \mathbf{H}^* \quad (5)$$

It is clear that at high SNR the MMSE decoder converges to the ZF decoder, while at low SNR the MMSE decoder prevents noise amplification by improving small eigen values before matrix inversion. Hence, the MMSE decoder reduces noise enhancement at the expense of complete interference cancellation. Since the transmitted symbols are drawn from a specific constellation with certain alphabet, a slicing operation is required as a post-processing operation for both the zero forcing and the MMSE decoders over the transmitted constellation. Successive interference cancellation receivers such as the Bell Laboratories Layered Space-Time (BLAST) receivers are among the suboptimal categories for solving (2) [6], [7]. Although the BLAST receiver, in its two forms; the diagonal (D-BLAST) and the vertical (V-BLAST) versions, outperform the MMSE and ZF decoders, they suffer from error propagation due to its successive nature. A near optimal receiver is the Sphere Decoder (SD) which finds the nearest lattice point inside a hypersphere, with variable radius, centered at the received signal point (2) [3], [8]. The SD has the best performance among all previous receivers but its performance varies as a function of the SNR [9]. A reminiscent of the sphere decoder is the Cube Decoder (CD) which finds the nearest lattice point inside a hypercube. The

Manuscript received Jun 30, 2009; revised November 30, 2009. The material in this paper was presented in part at the 17th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2009), Split-Hvar-Korula, Croatia, Sept. 2009.

M. Alghoniemy is with the Department of Electrical Engineering, University of Alexandria, Egypt (e-mail: alghoniemy@yahoo.com).

performance of the CD is inferior to the SD but it does not depend on the operating SNR [10]. For illustration purpose, the performance of different MIMO decoders for 4×4 MIMO system is illustrated in fig. 1. As can be seen from the figure, the performance of the MMSE receiver approaches the ZF at high SNR, as expected.

It should be mentioned that in this paper, we are only interested in the MMSE decoder and its connection to Tikhonov regularization. In particular, we will introduce a new decoder, the regularized decoder, which is considered an improved version of the MMSE decoder that takes into account the structure of the modulating constellation to gain performance improvement. In the next section, we will give a brief review on Tikhonov regularization and its connection to the MMSE receiver [11], [12]. In section 3, the regularized decoder is proposed and the penalty functions are introduced. Simulations are given in section 4.

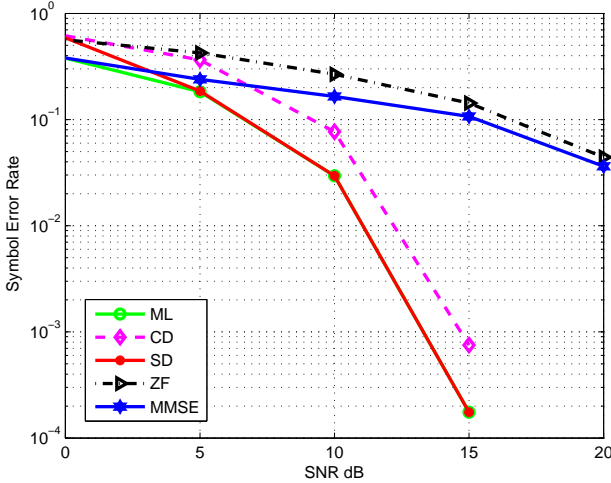


Fig. 1. Performance of MIMO decoders for 4×4 MIMO QPSK.

II. TIKHONOV REGULARIZATION

Regularization is a scalarization method that solves a multi-criterion optimization by converting the multi-criterion objective function into a positive weighted sum of the individual objectives. A special type of regularization is Tikhonov regularization solves a least squares problem while penalizing the squared norm of a function of the solution vector. In particular, Tikhonov regularization solves the following minimization problem [13], [14]

$$\mathbf{x}_{\text{tik}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \sqrt{\rho/m} \mathbf{H} \mathbf{x}\|^2 + \lambda \|\mathbf{L} \mathbf{x}\|^2 \quad (6)$$

where $\lambda > 0$ is the regularization parameter which controls how much weight is geared towards penalizing the norm of the solution vector. It should be noted that as $\lambda \rightarrow 0$, $\mathbf{x}_{\text{tik}} \rightarrow \mathbf{x}_{\text{zf}}$, and as $\lambda \rightarrow \infty$, \mathbf{x}_{tik} approaches the matched filter solution. \mathbf{L} is the regularization operator that usually approximates a high pass filter or a derivative operator as in image processing; Hence, $\|\mathbf{L} \mathbf{x}\|$ is a measure of the smoothness of the solution [15]. Thus the regularization problem (6) tries to solve the

least squares problem and at the same time penalizes a certain variation in the solution vector in the Euclidean norm sense. The above minimization problem can be re-written as an unconstrained least squares problem

$$\mathbf{x}_{\text{tik}} = \arg \min_{\mathbf{x}} \|\hat{\mathbf{y}} - \hat{\mathbf{H}} \mathbf{x}\|^2 \quad (7)$$

where $\hat{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}$ and $\hat{\mathbf{H}} = \begin{bmatrix} \sqrt{\rho/m} \mathbf{H} \\ \sqrt{\lambda} \mathbf{L} \end{bmatrix}$ where $\mathbf{0}$ is the zero vector of appropriate length. The problem in (7) has the following analytical solution which can be derived using the orthogonality principle [4]

$$\mathbf{x}_{\text{tik}} = \sqrt{\frac{\rho}{m}} \left(\frac{\rho}{m} \mathbf{H}^* \mathbf{H} + \lambda \mathbf{L}^* \mathbf{L} \right)^{-1} \mathbf{H}^* \mathbf{y} \quad (8)$$

which can be expressed as $\mathbf{x}_{\text{tik}} = \mathbf{G}_{\text{tik}} \mathbf{y}$ with

$$\mathbf{G}_{\text{tik}} = \sqrt{\frac{\rho}{m}} \left(\frac{\rho}{m} \mathbf{H}^* \mathbf{H} + \lambda \mathbf{L}^* \mathbf{L} \right)^{-1} \mathbf{H}^* \quad (9)$$

Consider the case where $\mathbf{L} = \mathbf{I}$ and $\lambda = 1$ in (8), we see that $\mathbf{x}_{\text{tik}} = \mathbf{x}_{\text{mmse}}$. Thus the MMSE decoder can be interpreted as a 2-norm regularized solution to the least squares problem. The previous interpretation coincides with the strategy that the MMSE decoder works, in the sense that it strikes a balance between interference cancellation "solving the least squares problem" and by limiting noise enhancement "reducing the norm of the solution vector".

III. REGULARIZED DECODERS

Given the previous interpretation of the MMSE decoder as a regularized least squares with penalty function equals to the ℓ_2 norm of the solution vector; it quadratically penalizes the elements of \mathbf{x} as they deviate from the origin which does not take into account the structure of the modulating constellation. In particular, assume that QPSK constellation is used in the modulation with elements equals to $\alpha(\pm 1 \pm j)$, where α is a normalizing factor, then the MMSE decoder quadratically penalizes the estimated symbols as they deviate from the origin, while it makes more sense to start penalizing them as they deviate away from $\alpha(\pm 1 \pm j)$.

Based on the previous observation, we propose a family of generalized decoders; which can be described as regularized least squares with different penalty functions that take into account the structure of the modulating constellation which leads to performance improvement. In particular, the generalized decoder solves the following regularization problem

$$\mathbf{x}_{\text{rgl}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \sqrt{\rho/m} \mathbf{H} \mathbf{x}\|^2 + \lambda \Phi(\mathbf{x}). \quad (10)$$

where $\Phi(\mathbf{x})$ describes a specific penalty function that penalizes the received symbols outside the convex hull of the constellation, these functions are convex and hence (10) is convex; which means that (10) can be solved efficiently with low complexity using convex optimization algorithms [14]. It should be noted that in the special case where $\Phi(\mathbf{x}) = \|\mathbf{x}\|^2$, the regularized decoder reduces to MMSE decoder. Similar to the ZF and MMSE, a slicing operation is required on \mathbf{x}_{rgl} in order to recover the estimated symbols.

A similar idea was proposed in [12], however, the proposed decoder here is considered more general in the sense that one can use different penalty functions for different constellations. It should be noted that the proposed decoder can also be considered as a generalization to the box-constrained least squares detector in the sense that the latter is a limiting case to the proposed decoder, this fact is illustrated in fig. 4 below [16].

For illustration purpose, we will consider two types of convex penalty functions, namely, the deadzone and infinity norm penalty functions.

A. Deadzone penalty

Two types of deadzone penalty functions are defined, deadzone-linear and deadzone-quadratic. In this case, $\Phi(\mathbf{x}) = \sum_i \phi_{dz}(x_i)$, $0 \leq i \leq m$. The deadzone-linear function $\phi_{dz-lin}(x)$ is defined as [14]

$$\phi_{dz-lin}(x) = \begin{cases} 0 & |x| \leq \alpha \\ \beta(|x| - \alpha) & |x| > \alpha \end{cases} \quad (11)$$

and the deadzone-quadratic function $\phi_{dz-quad}(x)$ is defined as

$$\phi_{dz-quad}(x) = \begin{cases} 0 & |x| \leq \alpha \\ \beta(x^2 - \alpha^2) & |x| > \alpha \end{cases} \quad (12)$$

where $\alpha \geq 0$ is the deadzone width and $\beta \geq 1$ determines the weight of the penalty function. In particular, fig. 2 shows a plot for the deadzone-linear and deadzone-quadratic functions with $\alpha = 1$ and two different β 's. It is clear that as β increases, the slope of $\phi_{dz}(x)$ increases and hence higher penalty is considered. It should be noted that if we set $\alpha = 0$ in the deadzone-linear (11), then it will be equivalent to the ℓ_1 norm penalty $\|\cdot\|_1$.

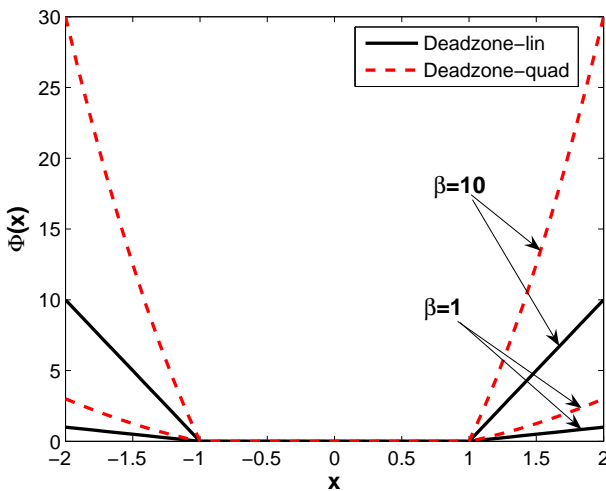


Fig. 2. Deadzone functions ($\alpha = 1$).

The advantage of using the deadzone penalty function over the usual ℓ_2 norm is that $\phi_{dz}(\mathbf{x})$ does not penalize the solution vector \mathbf{x} when it lies inside the deadzone region. Given that, if we choose α to be the maximum allowable symbol value

derived from the constellation, then \mathbf{x} will not be penalized when it lies inside the constellation. However, the elements of \mathbf{x} will be linearly or quadratically penalized, depending on using deadzone linear or deadzone quadratic functions, if they lie outside the constellation. Using the aforementioned choice of α , the solution tends to lie inside the constellation and better performance is gained.

B. Infinity norm penalty

Another possible penalty function is the infinity norm penalty. The ℓ_∞ norm penalizes the maximum absolute value of the estimated symbols in \mathbf{x} , which forces them not to deviate away from the constellation according to the penalty function $\Phi(\mathbf{x}) = \|\mathbf{x}\|_\infty$ [4].

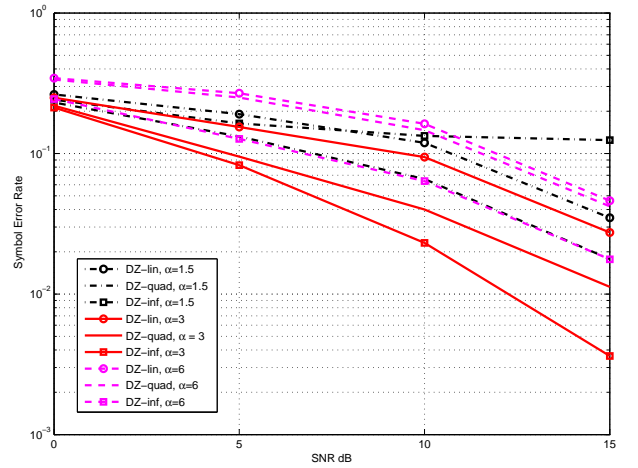


Fig. 3. Performance of 4×4 MIMO 16-QAM with different values of α .

IV. SIMULATION

In order to illustrate performance of the proposed decoders, we simulated a 2×2 and 4×4 MIMO system as in (1) and compared the performance of the ZF, MMSE and the regularized decoder. To solve problem (10) we used CVX, a package for specifying and solving convex programs [17], [18]. Fig. 3 shows the Symbol Error Rate (SER) as a function of the Signal to Noise Ratio (SNR) for the 4×4 MIMO system modulated using 16-QAM constellation $\alpha(\pm 1 \pm j)$ where $\alpha \in \{1, 3\}$, for different values of α , as it is clear from the figure, the best performance is obtained when $\alpha = 3$, the constellation size in this case.

The received constellation of the 4×4 MIMO system modulated using 16-QAM constellation is shown in fig. 4 after processing it with the dead-zone quadratic function with $\beta = 1$ and $\beta = 10$ and before slicing. It is clear that as β increases, the decoder heavily penalizes the received symbols outside the constellation, as expected. It should be noted that, the received constellation for high β behaves similar to the boxed constrained least squares (BCLS) receiver [16]. As is clear from the figure, the MMSE receiver has the least penalizing effect.

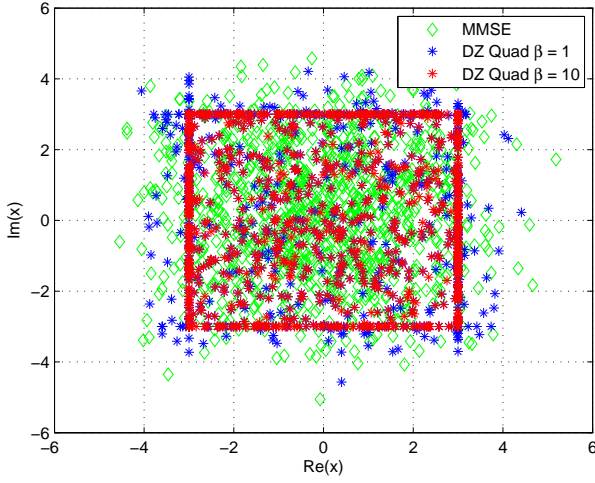
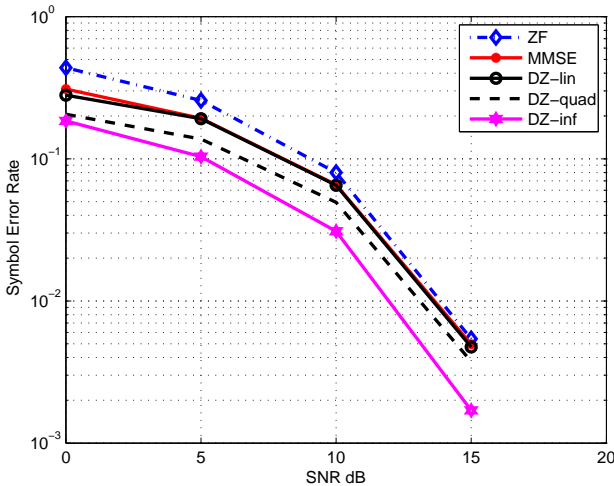
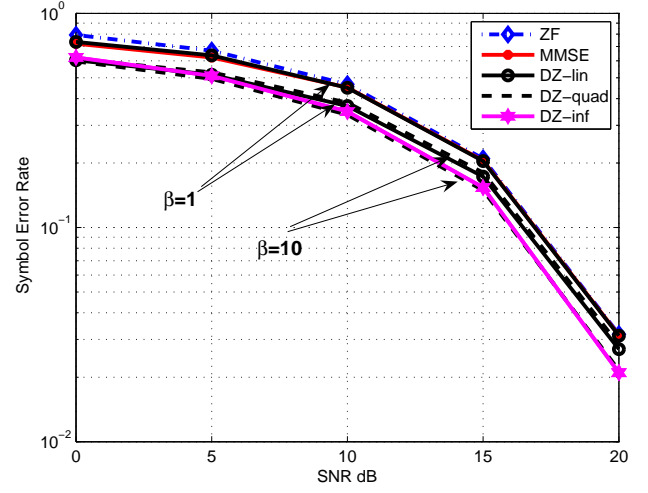
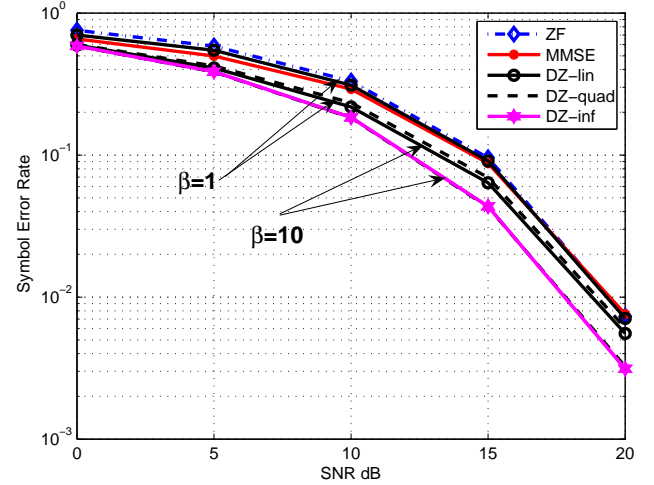
Fig. 4. Received constellation for 4×4 MIMO, 16-QAM.

Fig. 5 illustrates the behavior of the regularized decoder with $\beta = 1$, compared with the ZF and the MMSE decoders. We can see that the deadzone-quad decoder outperforms the MMSE decoder by 5-dB at low SNR and by a fraction of dBs at high SNR, while the ℓ_∞ decoder outperforms the MMSE decoder by a margin. It should be noted that the performance of the deadzone-lin decoder with $\beta = 1$ is comparable to the MMSE decoder, since for small x values, the quadratic and linear functions are comparable.

Finally, fig. 6 and fig. 7 show the performance of the proposed decoder as a function of the parameter β for 2×2 and 4×4 16-QAM system. In this case, it is clear that the performance of the regularized decoder with $\beta = 10$ is better than its performance with $\beta = 1$ as expected. It also should be noted that with high β , the deadzone-lin decoder outperforms the MMSE decoder due to the deviation between the linear and quadratic penalizing functions.

Fig. 5. Performance of the regularized decoder for QPSK 2×2 MIMO, $\beta = 1$.Fig. 6. Performance comparison for 2×2 MIMO 16-QAM.Fig. 7. Performance comparison for 4×4 MIMO, 16-QAM.

V. CONCLUSION

A family of regularized MIMO decoders has been proposed which is considered to be a generalization to the MMSE decoder. The MMSE decoder was shown to be equivalent to Tikhonov regularization, which is an ℓ_2 norm regularized least squares. The proposed decoders solve regularized least squares with different penalty functions that take into account the constellation structure by penalizing the received symbols outside the convex hull of the constellation. For illustration purpose, we have considered two penalty functions, the deadzone and ℓ_∞ norm penalty which led to performance improvement. It was shown that the regularized decoders outperform the MMSE at low as well as high SNR. Convex optimization routines can be used to implement the regularized decoders with low complexity.

REFERENCES

- [1] B. Natarajan, "Sparse Approximate Solutions to Linear Systems," *SIAM J. Comp.*, Vol. 24, pp. 227-234, Apr. 1995.
- [2] E. Agrell, T. Eriksson, A. Vardy, K. Zeger, "Closest Point Search in Lattices," *IEEE Transactions on Information Theory*, Vol. 48, No. 8, pp. 2201-2214, 2002.
- [3] Oussama Damen, Ammar Chkeif, and Jean-Claude Belfiore, "Lattice Code Decoder for Space-Time Codes," *IEEE Communications Letters*, VOL. 4, NO. 5, pp. 161-163, MAY 2000.
- [4] G. H. Goloub and C. F. V. Loan, *Matrix Computations*. The Johns Hopkins University Press, 1996.
- [5] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, May 2003.
- [6] 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.
- [7] 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," *Electron. Lett.*, vol. 35, pp. 14-16, Jan. 1999.
- [8] 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.
- [9] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm: Part I, the expected complexity," *IEEE Transactions on Signal Processing*, Vol. 53, No 8, pp. 2806-2818, Aug 2005.
- [10] Alghoniemy M; and Tewfik A. H. "Cube Decoding," *Proc. IEEE International Conf. on Communications (ICC)*. Vol. 3, pp. 1901-1905, May 2005.
- [11] Kasparis, C.; Piechocki, R.J.; Fletcher, P.N.; Nix, A.R. "A bootstrap multi-user detector for CDMA based on Tikhonov regularization," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing.*, Vol. 4, pp. 69-72, April 2003.
- [12] Piechocki, R.J. Kasparis, C. Nix, A.R. Fletcher, P.N. McGeehan, J.P. "Bootstrap frequency equalisation for MIMO wireless systems," *IEEE Global Telecommunications Conference, 2003. (GLOBECOM '03)*, Vol. 7, pp. 4175- 4179, Dec. 2003.
- [13] A. Neumaier, Solving ill-conditioned and singular linear systems: A tutorial on regularization, *SIAM Review*, vol. 40, no. 3, pp. 636-666, 1998.
- [14] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press 2004.
- [15] Per Christian Hansen, James G. Nagy, and Dianne P. OLeary, *Deblurring Images: Matrices, Spectra, and Filtering*, SIAM 2006.
- [16] Peng Hui Tan; Rasmussen, L.K.; Lim, T.J. , "Constrained maximum-likelihood detection in CDMA," *IEEE Trans. on Communications*, Vol. 49, No. 1, pp.142-153, Jan 2001.
- [17] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming (web page and software). <http://stanford.edu/~boyd/cvx>, February 2009.
- [18] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs, *Recent Advances in Learning and Control* (a tribute to M. Vidyasagar), V. Blondel, S. Boyd, and H. Kimura, editors, pages 95-110, *Lecture Notes in Control and Information Sciences*, Springer, 2008.



Masoud Alghoniemy received the 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 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 2001. Since April 2002, 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 806.16e. His research interests include wireless communications and sparse signal representation. 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. He is also listed in the 23rd edition of the Marquis Who's Who in the World.