A new family of MIMO receivers based on the equivalence between the MMSE and Tikhonov regularization

Masoud Alghoniemy
Electrical Engg. Dept., Faculty of Engg., Alexandria University, Alexandria Egypt
e-mail: *Alghoniemy@gmail*.com

In the Multi Input Multi Output antenna (MIMO) system, it is known that the Minimum Mean Squared Error (MMSE) receiver is equivalent to Tikhonov regularization. Given that, we developed a generalized receiver based on regularization with different penalty functions that take into account the structure of the modulating constellation, namely, dead zone and infinity norm penalty functions. Simulation results showed that the proposed receiver outperform the MMSE receiver by as high as 5-dB at low Signal Noise to Ratio (SNR) and 1-dB at high SNR.

فى نظام الهوائيات المتعددة المداخل والمتعددة المخارج، من المعروف أن المستقبل ذو الخطأ المتوسط المربع الأقل يكافىء تنظيم تيكونوف. بعد معرفة ذلك تقوم بتطوير مستقبل عام يعتمد على التنظيم باستخدام دوال معاقبة مختلفة والتى تأخذ فى الأعتبار شكل منظومة التعديل وهى دوال المعاقبة ذات المنطقة الميتة والدوال ذات المقياس اللانهائى. أعمال المحاكاة تبين أن المستقبل المقترح يتفوق على المستقبل ذو الخطأ المتوسط المربع الأقل بمقدار يصل الى ٥ ديسيبل فى حالة النسب العالية لمقدار الأشعة الى التشويش ويصل الى ١ ديسيبل فى حالة النسب المنخفضة لمقدار الأشعة الى التشويش.

Keywords: MIMO decoders, Tikhonov regularization

1. 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 (MIMO m×n), the relation between the transmitted and the received signal can be described as follows:

$$\mathbf{y} = \sqrt{\frac{\rho}{m}}\mathbf{Hx} + \mathbf{w}, \qquad (1)$$

where ρ is the expected value of the SNR at each receive antenna, $\mathbf{x}$ is the m×1 transmitted vector whose elements are complex symbols drown form the normalized Quadratic Amplitude Modulation (M-QAM) constellation with $E(\mathbf{xx}^T) = \mathbf{I}$, where M is the constellation order and $\mathbf{I}$ is the m × m identity matrix. $\mathbf{y}$ is the n×1 received vector, $\mathbf{H}$ is the n × m channel matrix with n≥m, whose elements represent the Independent Identically Distributed (IID) flat fading Complex Normal channel gains $h_{ij} \approx CN\,(0,\sigma_h^2)$. Without loss of generality, $\mathbf{w}$ is n × 1 IDD zero mean complex white Gaussian noise, uncorrelated with the transmitted symbols, with $w_i \approx CN(0,1)$. For the Gaussian noise scenario, the optimum decoder is the Maximum Likelihood (ML) decoder which finds the most likely input vector $\mathbf{x_{ml}}$ according to

$$\mathbf{x_{ml}} = \arg \min_{x\subset\Lambda}\left\|\mathbf{y} - \sqrt{\frac{\rho}{m}}\mathbf{Hx}\right\|^2 \qquad (2)$$

where Λ is the lattice whose points represent all possible combinations of $\mathbf{x}$. The problem eq. (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 discrete nature of the lattice Λ. Linear decoders have been used to obtain an approximate solution with low complexity, the simplest linear decoder is the Zero-Forcing (ZF) decoder in which the constraint $\mathbf{x} \subset \Lambda$ is relaxed and the domain in this case is R^n. The zero-forcing decoder inverts the channel in order to cancel spatial interference, in particular $\mathbf{x_{zf}} = \mathbf{G_{zf}}\,\mathbf{y}$ with

$$\mathbf{G_{zf}} = \sqrt{\frac{m}{\rho}}\left(\mathbf{H}^*\mathbf{H}\right)^{-1}\mathbf{H}^* \quad , \qquad (3)$$

is the pseudo-inverse of the channel [2]. 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 would amplify the contaminating noise. In an effort to reduce noise enhancement, the Minimum Mean Squared Error (MMSE) decoder is used to strike a balance between interference cancellation and noise enhancement [3]. The MMSE decoder finds the solution $\mathbf{x_{mmse}} = \mathbf{G_{mmse}}\,\mathbf{y}$, where

$$\mathbf{x_{mmse}} = \arg\ \mathbf{min}\ \mathrm{E}\|\mathbf{Gy}-\mathbf{x}\|^2, \tag{4}$$

which has the analytical solution

$$\mathbf{G_{mmse}} = \sqrt{\frac{\rho}{m}}\left(\frac{\rho}{m}\mathbf{H}^*\mathbf{H}+\mathbf{I}\right)^{-1}\mathbf{H}^*. \tag{5}$$

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. A near optimal receiver is the sphere decoder which finds the nearest lattice point inside a sphere centred at the received signal point eq. (2) [4]. Successive interference cancellation receivers such as the D-BLAST and the V-BLAST are among the suboptimal categories for solving eq. (2) [5, 6]. In this paper, we are only interested in the MMSE decoder and its connection to Tikhonov regularization.

2. Tikhonov regularization

In this section we review the known results that the MMSE receiver is equivalent to Tikhonov regularization [7].

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 objectives. In particular, Tikhonov regularization solves a least squares problem while penalizing the squared norm of a function of the solution vector [9, 10]. The penalizing parameter is called the regularization parameter; Tikhonov regularization solves the following minimization problem

$$\mathbf{x_{tik}} = \arg\ \mathbf{min}\left\|\mathbf{y}-\sqrt{\frac{\rho}{m}}\mathbf{Hx}\right\|^2 + \lambda\|\mathbf{Lx}\|^2, \tag{6}$$

towards penalizing the norm of the solution vector. It should be noted that as $\lambda \rightarrow 0$, $\mathbf{x_{tik}} \rightarrow \mathbf{x_{zf}}$, and as $\lambda \rightarrow \infty$, $\mathbf{x_{tik}}$ approaches the matched filter solution. $\mathbf{L}$ is an operator that usually approximates a high pass filter or a derivative operator as in image processing; hence, $\|\mathbf{Lx}\|$ is a measure of the smoothness of the solution [11].

Thus the regularization problem eq. (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

$$\mathbf{x_{tik}} = \arg\ \underset{\mathbf{x}}{\mathbf{min}}\|\mathbf{y}'-\mathbf{H}'\mathbf{x}\|^2. \tag{7}$$

Where $\mathbf{y}' = \begin{bmatrix}\mathbf{y}\\ \mathbf{0}\end{bmatrix}$ and $\mathbf{H}' = \begin{bmatrix}\sqrt{\frac{\rho}{m}}\mathbf{H}\\ \sqrt{\lambda}\mathbf{L}\end{bmatrix}$, $\mathbf{0}$ is the zero vector of appropriate length. The problem in eq. (7) has the following analytical solution which can be derived using the orthogonality principle [2].

$$\mathbf{x_{tik}} = \sqrt{\frac{\rho}{m}}\left(\frac{\rho}{m}\mathbf{H}^*\mathbf{H}+\lambda\mathbf{L}^*\mathbf{L}\right)^{-1}\mathbf{H}^*\mathbf{y}, \tag{8}$$

which can be expressed as $\mathbf{x_{tik}} = \mathbf{G_{tik}}\,\mathbf{y}$ with

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

Consider the case where **L = I** and $\lambda = 1$ in eq. (8), we see that $\mathbf{x_{tik}} = \mathbf{x_{mmse}}$.

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".

3. 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 **x** as they deviate from the origin which does not take into account the structure of the constellation. In particular, assume that Quadrature Phase Shift Keying (QPSK) constellation is used in the modulation with elements equals to ± 1 ± j, 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 ± 1 ± j.

Based on the previous observation, we propose a generalized decoder as regularized least squares with other 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

$$\mathbf{x_{rgl}} = \arg\ \mathbf{min}\left\| \mathbf{y} - \sqrt{\frac{\rho}{m}}\mathbf{Hx} \right\|^2 + \lambda \mathbf{\Phi}(\mathbf{x}), \qquad (10)$$

where $\mathbf{\Phi}(\mathbf{x})$ describes a specific penalty function as shown below, these functions are convex and hence eq. (10) is convex. It should be noted that in the special case where $\mathbf{\Phi}(\mathbf{x}) = \|\mathbf{x}\|^2$, 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.

It should be noted that the proposed detector is considered as a generalization to the box-constrained least squares detector in the sense that the box-constrained least squares detector is a limiting case for the proposed detector [8].

3.1. Deadzone penalty

Two types of deadzone penalty functions are defined, deadzone-linear and deadzone-quadratic. In this case,

$$\mathbf{\Phi}(\mathbf{x}) = \sum_i \varphi_{dz}(x_i) \qquad 0 \le \mathrm{i} \le \mathrm{m}\,.$$

The deadzone-linear function $\phi_{dz-lin}(x)$ is defined as [10]

$$\phi_{dz-lin}(x) = \begin{cases} 0 & |x| \le \alpha \\ \beta\left(|x| - \alpha\right) & |x| > \alpha, \end{cases} \qquad (11)$$

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

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

where $\alpha \ge 0$ is the deadzone width and $\beta \ge 1$ determines the weight of the penalty function. In particular, fig. 1 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 eq. (11), then it will be equivalent to the ℓ_1 norm penalty $\|\ \|_1$.

The advantage of using the deadzone penalty function over the usual ℓ_2 norm is that $\phi_{dz}(x)$ does not penalize the solution vector **x** when it lies inside the deadzone region.

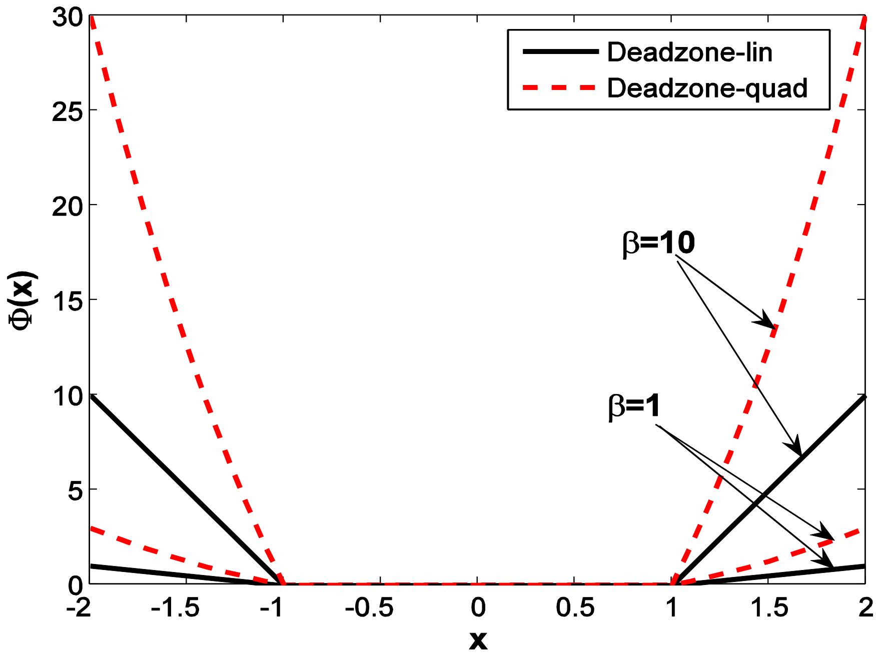

Fig. 1. Deadzone functions α =1.

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.

3.2. 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 $\mathbf{\Phi}(\mathbf{x}) = \|\mathbf{x}\|_\infty$ [2].

4. Simulation

In order to illustrate performance of the proposed decoders, we simulated a 2 × 2 and 4 × 4 MIMO system as in eq. (1) and compared the performance of the ZF, MMSE and the regularized decoder. To solve problem eq. (10) we used CVX, a package for specifying and solving convex programs [12].

In particular, fig. 2 illustrates the behaviour of the regularized decoder with β=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 β =1 is comparable to the MMSE decoder, since for small x values, the quadratic and linear functions are comparable. On the other hand, fig. 3 and 4 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 β = 10 is better than its performance with β = 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 function.

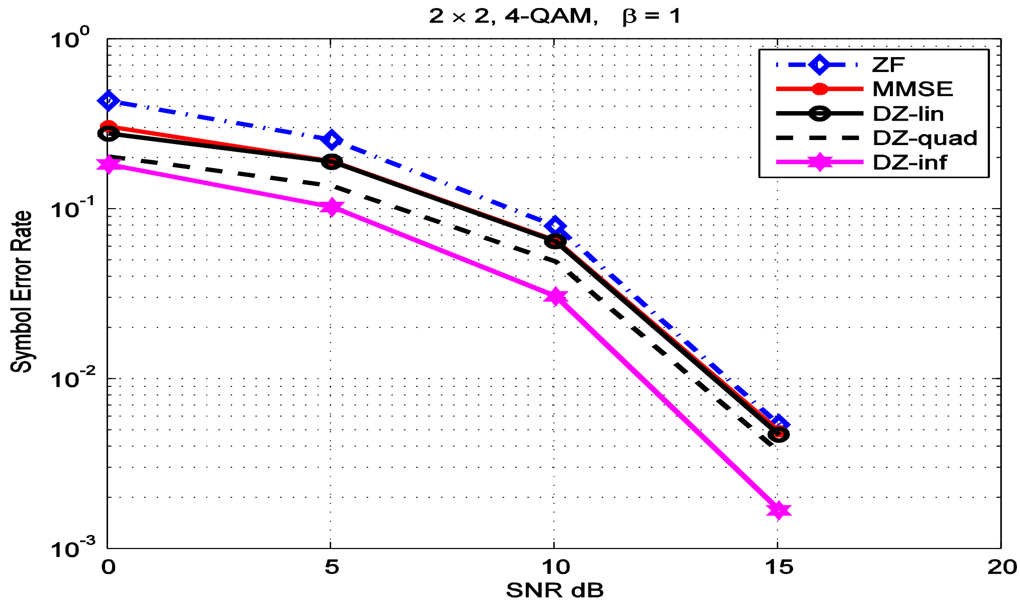

Fig. 2. Performance for 2 × 2 MIMO QPSK, β =1.

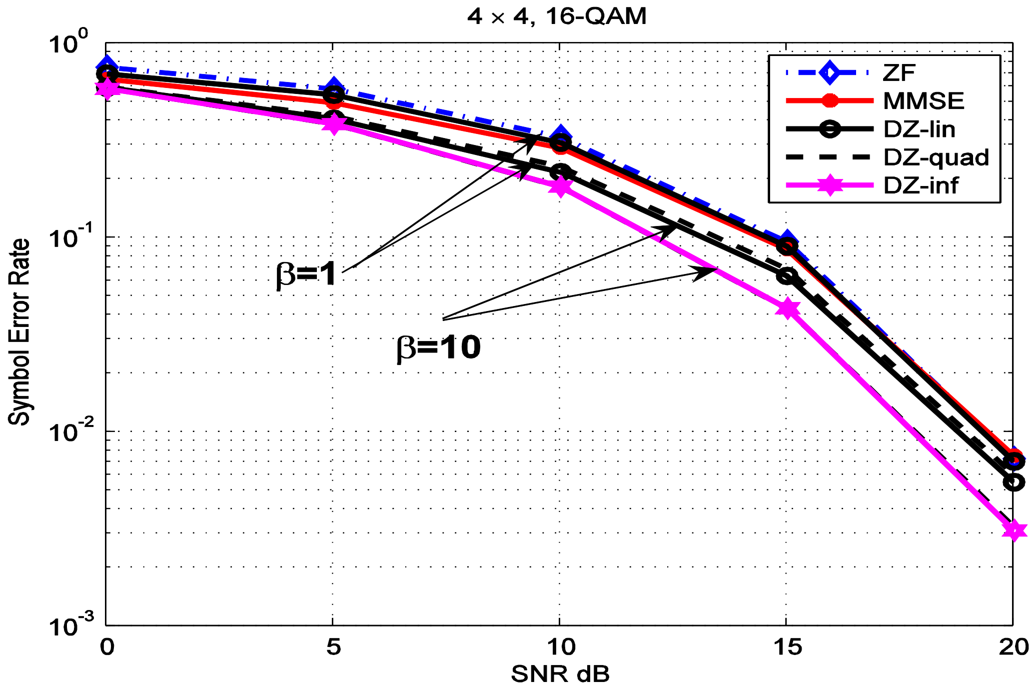

Fig. 3. Performance comparison for 2 × 2 MIMO, 16-QAM.

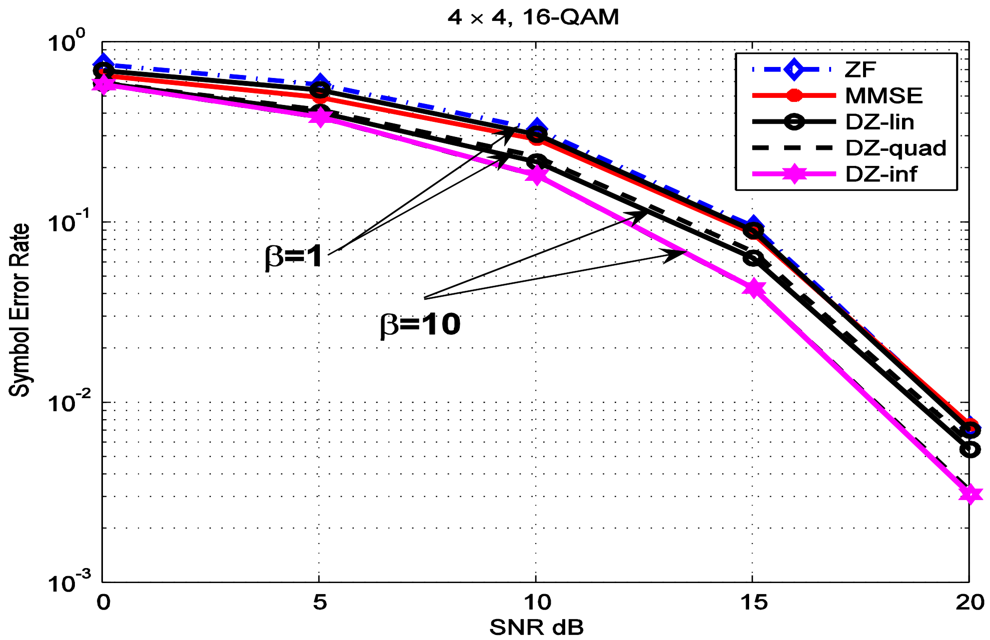

Fig. 4. Performance comparison for 4 × 4 MIMO, 16-QAM.

5. Conclusions

A new MIMO decoder 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 a ℓ_2 norm regularized least squares. The proposed decoder solves regularized least squares with different penalty functions that take into account the constellation structure, namely deadzone and ℓ_∞ norm penalty which lead to performance improvement. It was shown that the regularized decoder outperforms the MMSE for low as well as high SNR. The proposed decoder has the same complexity as the MMSE.

References

[1] B. Natarajan, "Sparse Approximate Solutions to Linear Systems," SIAM J. Comp., Vol. 24, pp. 227-234 (1995).

[2] G.H. Goloub and C.F.V. Loan, "Matrix Computations", the Johns Hopkins University Press (1996).

[3] A. Paulraj, R. Nabar and D. Gore, "Introduction to Space-Time Wireless Communications", Cambridge University Press (2003).

[4] 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 (1985).

[5] 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 (2), pp. 41–59 (1996).

[6] G.D. Golden, G.J. Foschini, R.A. Valenzuela and P.W. Wolniansky, "Detection Algorithm and Initial Laboratory Results Using V-BLAST Space–Time Communication Architecture," Electron. Lett., Vol. 35, pp. 14–16 (1999).

[7] C. Kasparis, R.J. Piechocki, P.N. Fletcher and A.R. Nix, "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, (2003).

[8] Peng Hui Tan, L.K. Rasmussen and T.J. Lim, "Constrained Maximum-Likelihood Detection in CDMA," IEEE Trans. on Communications, Vol. 49 (1), pp. 142-153, Jan 2001.

[9] A. Neumaier, "Solving Ill-Conditioned and Singular Linear Systems: A Tutorial on Regularization", SIAM Review, Vol. 40 (3), pp. 636-666 (1998).

[10] Stephen Boyd and Lieven Vandenberghe, Convex Optimization, Cambridge University Press (2004).

[11] Per Christian Hansen, James G. Nagy, and Dianne P. O'Leary, Deblurring Images: Matrices, Spectra, and Filtering, SIAM (2006).

[12] M. Grant and S. Boyd. CVX, Matlab Software for Disciplined Convex Programming (web page and software). http://stanford.edu/~boyd/cvx (2009).

Received July 6, 2009
Accepted July 25, 2009