

A Trust–Region Approach to the Regularization of Large–Scale Discrete Ill–Posed Problems

Marielba Rojas ^{*} Danny C. Sorensen [†]

January 12, 2000

**Technical Report 99–26, Department of Computational and Applied Mathematics,
Rice University, Houston.**

Abstract

We consider the solution of large–scale least squares problems where the coefficient matrix comes from the discretization of an ill–posed operator and the right–hand side contains noise. Special techniques known as regularization methods are needed to treat these problems in order to control the effect of the noise on the solution. We pose the regularization problem as a trust–region subproblem and solve it by means of a recently developed method for the large–scale trust–region subproblem. We present numerical results on test problems, an inverse interpolation problem with real data, and a model seismic inversion problem with real data.

AMS classification: Primary: 86A22. Secondary: 65K10, 90C06

Key words and phrases: Regularization, constrained quadratic optimization, trust region, Lanczos method, ill–posed problems, inverse problems, seismic inversion.

1 Introduction

Discrete forms of ill–posed problems arise when we discretize the continuous operator from an ill–posed problem and introduce experimental data contaminated by noise.

^{*}Department of Computational and Applied Mathematics, Rice University, 6100 Main St., Houston, TX 77005-1892, USA (mrojas@caam.rice.edu). This author was supported by NSF cooperative agreement CCR-9120008.

[†]Department of Computational and Applied Mathematics, Rice University, 6100 Main St., Houston, TX 77005-1892, USA (sorensen@caam.rice.edu). This author was supported in part by NSF cooperative agreement CCR-9120008, and by ARPA contract number DAAL03-91-C-0047 (administered by the U.S. Army Research Office).

One of the main sources of ill-posed problems are inverse problems, where we want to determine the internal structure of a system from the response of the system to external stimuli. Inverse problems arise in many important applications such as image processing, signal processing, seismic inversion, and medical and seismic tomography. Discrete forms of ill-posed problems are usually formulated as linear systems or least squares problems. The focus of this paper is the numerical treatment of large-scale discrete ill-posed least squares problems.

We consider the problem

$$\min \|Ax - \bar{b}\| \tag{1}$$

with $A \in \mathbb{R}^{m \times n}$ and $\bar{b} \in \mathbb{R}^m$, such that A comes from the discretization of a continuous ill-posed operator and $\bar{b} = b + s$, where b is an *exact, unknown* data vector and s is a random vector of uncorrelated noise. The norm is the Euclidean norm throughout the paper.

We will assume that the matrix A in problem (1) is large and might not be available explicitly but the action of A and A^T on vectors of the appropriate dimensions is known. We will also assume that errors in A , due to discretization or finite-precision representation, are small in comparison to the noise in the right-hand side. Finally, we will not assume any particular structure for A .

If a reasonably accurate discretization is used, the matrix A will be highly ill-conditioned with a singular spectrum that decays to zero gradually, a large cluster of small singular values, and high-frequency components of the singular vectors associated with small singular values. If, in addition, the Discrete Picard Condition [9] holds, we have that components of the exact data vector b in the direction of the left singular vectors, decay to zero faster than the singular values of A , while components of the noise vector s remain constant. Therefore, components corresponding to small singular values are magnified by the noise.

As a consequence of the ill conditioning of the matrix A and the presence of noise in the right-hand side, standard methods on problem (1) produce meaningless solutions with extremely large norm. Therefore, to solve these problems, we need to use special techniques known as *regularization* or *smoothing methods*. These methods aim to recover information about the desired solution of the unknown problem with exact data from the solution of a better conditioned problem that is related to the problem with noisy data but incorporates additional information about the desired solution. The formulation of this new problem involves a special parameter known as the *regularization parameter*, used to control the effect of the noise on the solution. Excellent surveys on regularization methods can be found for example in [8], [11] and more recently in [19].

While there are many alternatives for small to medium-scale problems, large-scale problems remain a challenge. However, in recent years some interesting methods have been proposed, such as the method of Golub and von Matt [7], the

method of Björck, Grimme and van Dooren [3], the method of Calvetti, Reichel and Zhang [4], as well as several variants of the Conjugate Gradient Method on the Normal Equations, including the use of preconditioners chosen according to the structure of the problem, [13], [15] and [18].

The most common approach in real applications is to apply the Conjugate Gradient Method to the Normal Equations, relying on observed intrinsic regularization properties of the method. However, this approach lacks a systematic termination criterion and as a consequence, the approximate solution must be visually inspected at each iteration to determine when to stop.

In this paper we apply a recently developed method [22] for the large-scale trust-region subproblem to the regularization of discrete forms of ill-posed problems from a variety of applications. The method relies on matrix-vector products only, has low and fixed storage requirements, robust stopping criteria and computes both a solution and the corresponding Tikhonov regularization parameter. Most of the results presented here are based on the work in [21]. The organization of the paper is the following. In Section 2 we describe our regularization approach and show its connection with Tikhonov regularization and with the trust-region subproblem. In Section 3 we describe the structure of the trust-region subproblem and show its special properties in the discrete ill-posed case. In Section 4 we describe the method LSTRS from [22] and discuss the issues related to ill-posed problems. In Section 5 we present numerical results of LSTRS on regularization problems, including test problems from the Regularization Tools package [10], an inverse interpolation problem with real data, and a model seismic inversion problem. We present some conclusions in Section 6.

2 Regularization through Trust Regions

As we mentioned before, regularization involves the formulation of an auxiliary problem related to both the original problem with exact data and the problem with noisy data. One of the most popular regularization approaches is the classical Tikhonov regularization approach [25] in which the auxiliary problem is the following

$$\min \|Ax - \bar{b}\|^2 + \varepsilon^2 \|Lx\|^2, \quad (2)$$

where $\varepsilon^2 > 0$ is the regularization parameter. This parameter acts as a penalty parameter on either the norm of the solution when $L = I$, or the smoothness of the solution when $L \neq I$ is a discrete form of first derivative. If L is nonsingular then a change of variables yields an equivalent problem with $L = I$. We will assume this is the case throughout the paper.

Observe that for a given ε , problem (2) becomes a damped least squares problem that we can solve with standard numerical linear algebra techniques for medium

and large-scale problems (cf. [6]). However, determining an optimal value for the Tikhonov regularization parameter is usually as difficult as the original problem and most of the methods currently available require the solution of several problems of type (2) for different values of the parameter. An exception to this situation is the method of Calvetti, Reichel and Zhang [4] where there is an elegant way of computing the parameter given the noise level in the data.

In this work we use an equivalent formulation of problem (2), as a quadratically constrained least squares problem

$$\begin{aligned} \min \quad & \|Ax - \bar{b}\| \\ \text{s.t.} \quad & \|x\| \leq \Delta \end{aligned} \tag{3}$$

where $\Delta > 0$. Assuming $\bar{b} \notin \mathcal{R}(A)$, any feasible point is a regular point and therefore the Karush–Kuhn–Tucker conditions for a minimizer are $(A^T A - \lambda_* I) x_* = -A^T \bar{b}$, $\lambda_* \leq 0$, and $\lambda_*(\|x_*\| - \Delta) = 0$. And observe that since (3) is a convex quadratic problem, these conditions are both necessary and sufficient. Equivalence with problem (2) follows directly by setting the optimal value of the Tikhonov regularization parameter equal to minus the optimal Lagrange multiplier for the quadratically constrained least squares problem, i.e. $\varepsilon_*^2 = -\lambda_*$.

While Tikhonov regularization involves the computation of a parameter with no *a priori* physical meaning, the quadratically constrained least squares formulation has the advantage that, in some applications, the physical properties of the problem determine an optimal value for the norm constraint Δ . This is the case, for example, in image restoration where Δ represents the energy of the target image.

An additional advantage of the quadratically constrained least squares formulation is that (3) is a special case of a well-known problem in optimization, namely, the problem of minimizing a quadratic on a sphere or trust-region subproblem

$$\begin{aligned} \min \quad & \frac{1}{2}x^T Hx + g^T x \\ \text{s.t.} \quad & \|x\| \leq \Delta \end{aligned} \tag{4}$$

where $H \in \mathbb{R}^{n \times n}$, $g \in \mathbb{R}^n$ and $\Delta > 0$. Problem (3) is a special case of (4) when $H = A^T A$ and $g = -A^T \bar{b}$.

The high degree of structure of the trust-region subproblem leads to strong theoretical properties and makes it possible to design efficient solution methods. Therefore we shall formulate the regularization problem as a trust-region subproblem.

3 The Trust–Region Subproblem

3.1 Structure of Problem

A first observation about the trust–region subproblem is that it always has a solution. A not so obvious and quite remarkable fact about the problem is the existence of a characterization of its solutions. The result is contained in the following Lemma from [23].

Lemma 3.1 ([23]) *A feasible vector x_* is a solution to (4) with corresponding Lagrange multiplier λ_* if and only if x_*, λ_* satisfy $(H - \lambda_* I)x_* = -g$ with $H - \lambda_* I$ positive semidefinite, $\lambda_* \leq 0$ and $\lambda_*(\Delta - \|x_*\|) = 0$.*

Proof. See [23]. □

The optimality conditions imply that all the solutions of the trust–region subproblem are of the form $x = -(H - \lambda I)^\dagger g + z$ for $z \in \mathcal{N}(H - \lambda I)$. These solutions may lie in the interior or on the boundary of the feasible set. There are no solutions on the boundary if and only if H is positive definite and $\|H^{-1}g\| < \Delta$ (see [17]). In this case, the unique interior solution is $x = -H^{-1}g$ with Lagrange multiplier $\lambda = 0$. Boundary solutions satisfy $\|x\| = \Delta$ with $\lambda \leq \delta_1$, where δ_1 is the smallest eigenvalue of H . The case $\lambda = \delta_1$ can only occur if $g \perp \mathcal{S}_1$ where $\mathcal{S}_1 \equiv \mathcal{N}(H - \delta_1 I)$ and $\|(H - \delta_1 I)^\dagger g\| \leq \Delta$. This corresponds to the so-called *hard case*, which poses great difficulties for the numerical solution of the trust–region subproblem since in this case it is necessary to compute an approximate eigenvector associated with the smallest eigenvalue of H . Moreover, in practice g will be nearly orthogonal to \mathcal{S}_1 and we can expect greater numerical problems in this case. We call this situation a *near hard case*, and since whenever $g \perp \mathcal{S}_1$ there is the possibility for the hard case to occur, we call this a *potential hard case*. We show in §3.2 that the potential hard case and near hard case are precisely the common cases for discrete ill–posed problems.

The conditions in Lemma 3.1 are computationally attractive since they provide a means for reducing the problem of computing boundary solutions for the trust–region subproblem from an n –dimensional problem to a zero–finding problem in one variable. We can accomplish this for example, by solving the *secular equation*

$$\Delta - \|x\| = 0 \tag{5}$$

monitoring λ so $H - \lambda I$ is positive semidefinite. Assuming that we can compute the Cholesky factorization of $H - \lambda I$ then Newton’s method is particularly efficient for solving an equation equivalent to (5). This approach, due to Moré and Sorensen [17], is the method of choice whenever it is possible to compute the Cholesky factorization of matrices of the form $H - \lambda I$. However, in some applications this computation

may be prohibitive either because of storage considerations or because the matrix H is not explicitly available. Therefore, we need other strategies to treat the problem in those cases.

3.2 Discrete Ill-Posed Trust-Region Subproblem

We now study the trust-region subproblem in the special case when $H = A^T A$ and $g = -A^T \bar{b}$, where A comes from the discretization of a continuous ill-posed operator and \bar{b} contains noise. We will show that the potential hard case is the common case for these problems and also that it will occur in a *multiple* instance, with g orthogonal to the eigenpaces associated with several of the smallest eigenvalues of H . This follows from the following result.

Lemma 3.2 *Let $H = A^T A$ and $g = -A^T \bar{b}$, with $\bar{b} = b + s$. Suppose σ_k is the k th largest singular value of A with multiplicity m_k . Suppose u_j, v_j , $1 \leq j \leq m_k$ are left and right singular vectors associated with σ_k . Then*

$$g^T v_j = -\sigma_k (u_j^T b + u_j^T s) .$$

Proof. The result follows directly assuming $A = U \Sigma V^T$ is a singular value decomposition of A , since this yields $g = -V \Sigma U^T \bar{b}$ with V orthogonal. \square

Lemma 3.2 implies that if σ_k is very small then g will be nearly orthogonal to the subspace spanned by the right singular vectors associated with σ_k , regardless of the noise level in \bar{b} . For large noise level and σ_k not so small, g will not be nearly orthogonal to that space. In discrete ill-posed problems, the matrix A has a large cluster of very small singular values and therefore we can expect g to be nearly orthogonal to the right singular vectors associated with such singular values. Since these vectors are eigenvectors corresponding to the smallest eigenvalues of $A^T A$, this means that g will be orthogonal to the eigenspaces corresponding to several of the smallest eigenvalues of $A^T A$ and that the potential hard case will occur in a *multiple* instance. Figure 1 illustrates this situation for problem **foxgood** from the Regularization Tools package by Hansen [10]. The problem is of dimension 300 and in the logarithmic plot we observe that $g^T v_k$ is of order 10^{-15} for approximately 292 eigenvectors.

4 LSTRS for Discrete Ill-Posed Problems

In this section we give a brief description of the method LSTRS from [22] and discuss the advantages of using this method for solving large-scale discrete ill-posed trust-region subproblems. LSTRS is based on formulating the trust-region subproblem as

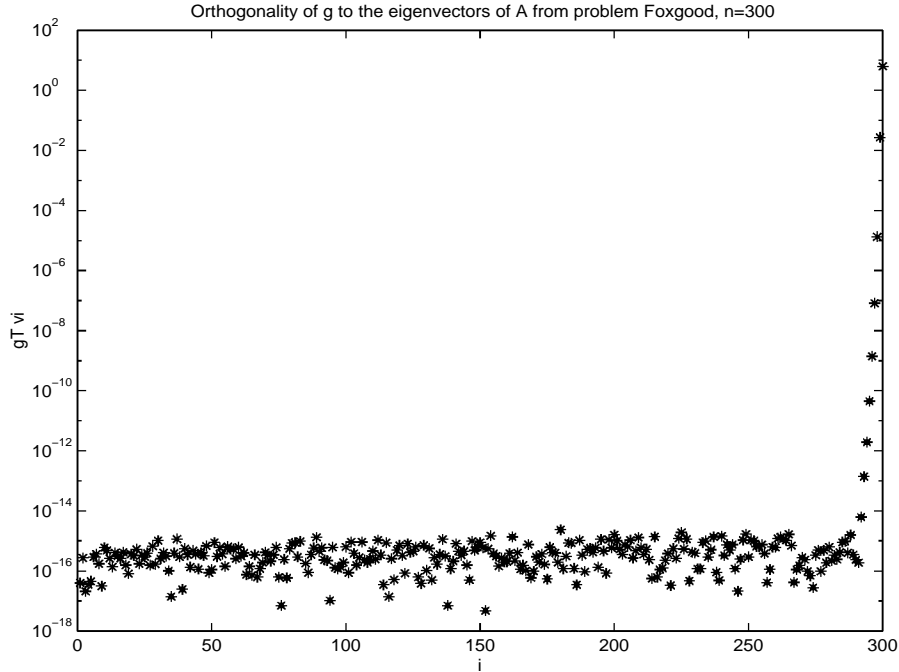


Figure 1: Orthogonality of g with respect to the eigenvectors of a discretized ill-posed operator.

a parameterized eigenvalue problem. Such formulation comes from the observation that for $\alpha = \lambda_* - g^T x_*$, problem (4) is equivalent to

$$\begin{aligned} \min \quad & \frac{1}{2} y^T B_\alpha y \\ \text{s.t.} \quad & y^T y \leq 1 + \Delta^2, \quad e_1^T y = 1, \end{aligned} \quad (6)$$

where e_1 is the first canonical unit vector in \mathbb{R}^{n+1} and $B_\alpha = \begin{pmatrix} \alpha & g^T \\ g & H \end{pmatrix}$. The solution of the trust-region subproblem consists of the last n components of the solution of problem (6).

Problem (6) suggests that if we know the optimal value for α , we can solve the trust-region subproblem by solving an eigenvalue problem for the smallest eigenvalue of B_α and an eigenvector with special structure. To see this, observe that if $\{\lambda, (1, x^T)^T\}$ is an eigenpair of B_α then

$$\begin{pmatrix} \alpha & g^T \\ g & A \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \lambda$$

which is equivalent to

$$\alpha - \lambda = -g^T x \quad \text{and} \quad (H - \lambda I)x = -g. \quad (7)$$

If λ is the smallest eigenvalue of B_α and since the eigenvalues of H interlace the eigenvalues of B_α by Cauchy Interlace Theorem (cf. [20]), then $H - \lambda I$ is positive semidefinite. Therefore, we have that two of the optimality conditions in Lemma 3.1 are satisfied in this case. If in addition, $\lambda \leq 0$ and $\|x\| = \Delta$, we will have a solution for the trust-region subproblem.

LSTRS consists in iteratively adjusting the parameter α to drive it towards the optimal value $\alpha_* = \lambda_* - g^T x_*$. This is accomplished in the following way. Let $\phi(\lambda) = -g^T x$ for x satisfying $(H - \lambda I)x = -g$, and note that $\phi'(\lambda) = x^T x$. Both ϕ and ϕ' are rational functions with poles at a subset of the eigenvalues of H , and Figure 2 illustrates the typical behavior of $\phi(\lambda)$. The LSTRS iteration is based on approximately solving the secular equation (5), using rational interpolation on ϕ and ϕ' . Observe that, in view of (7), we can obtain convenient interpolation points by solving eigenvalue problems for the smallest eigenvalue of B_α , for different values of the parameter α . LSTRS computes the interpolation points in this way, using the Implicitly Restarted Lanczos Method (IRLM) [24] as implemented in ARPACK [16] to solve the eigenvalue problems. The IRLM has fixed storage requirements and relies upon matrix-vector products only, features that make it suitable for large-scale problems.

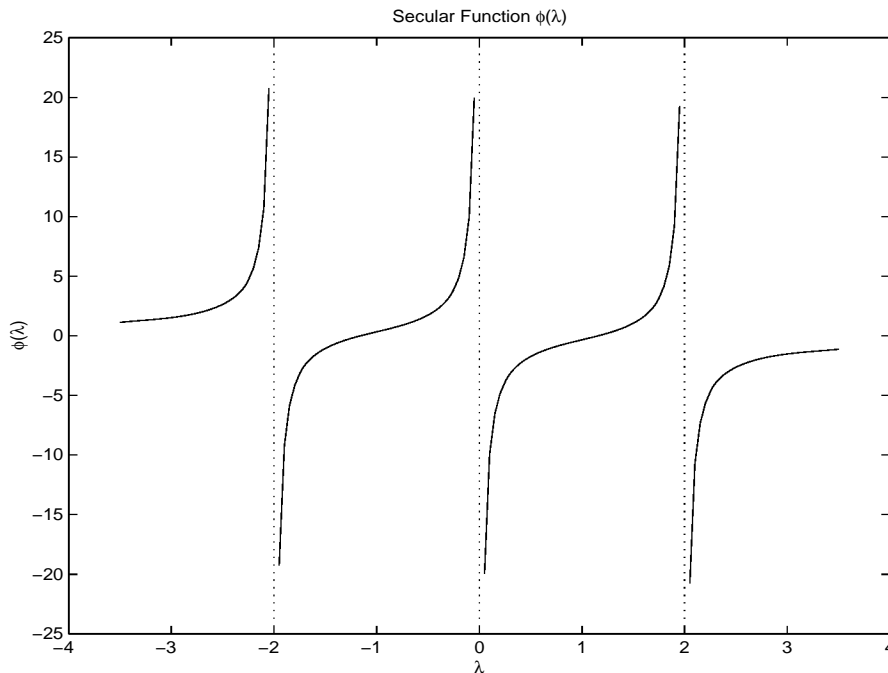


Figure 2: Function $\phi(\lambda)$.

The strategy described above works as long as the smallest eigenvalue of B_α has a corresponding eigenvector that can be safely normalized to have first component

one. The adjustment of the parameter becomes very difficult in the hard case and near hard case since in these situations the smallest eigenvalue of B_α might not have a corresponding eigenvector with the desired structure. Moreover, in the near hard case δ_1 is a weak pole of $\phi(\lambda)$ and the function becomes very steep around this value as we can observe in Figure 3. This makes the interpolation problem very ill-conditioned as we can observe in Figure 3. LSTRS relies on the complete characterization of the hard case given in [21] to proceed with the iteration even when the desired eigenvector cannot be normalized to have first component one, and to compute nearly optimal solutions in any instance of the hard case including *multiple* occurrences as in ill-posed problems. Figure 4 shows a plot of the function $\phi(\lambda)$ for a viscoacoustic wave equation, where we can observe the multiple occurrence of the (potential) near hard case.

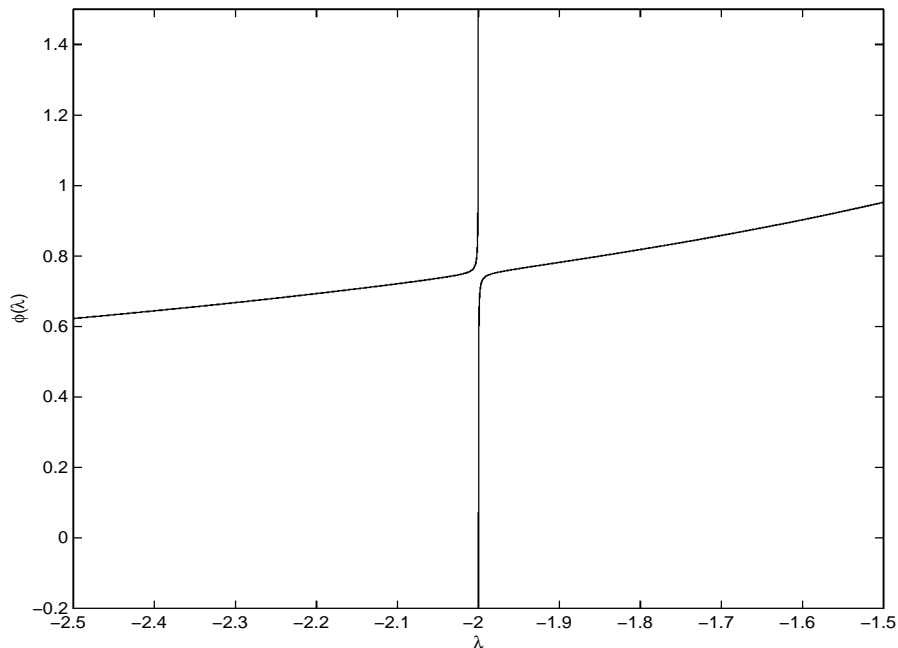
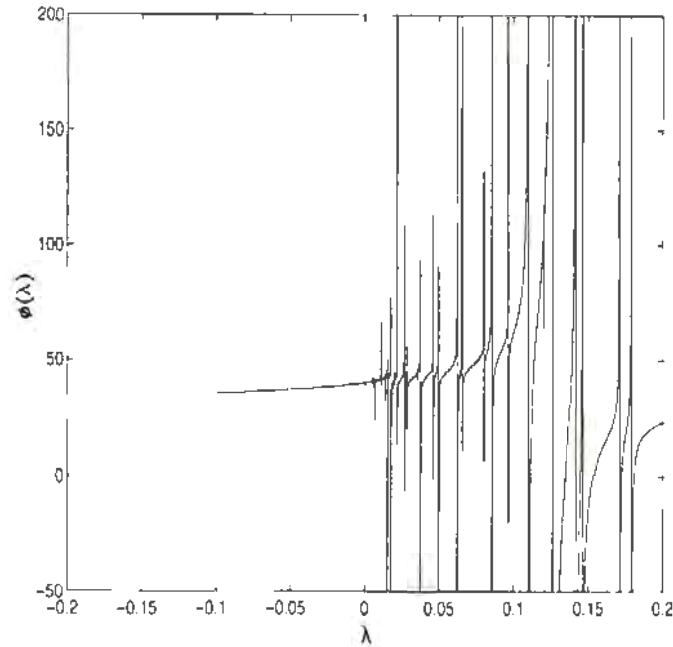


Figure 3: Function $\phi(\lambda)$ in the near hard case.

From the above presentation we see that LSTRS has desirable features for solving large-scale trust-region subproblems in general, and for handling discrete ill-posed problems in particular. This is not surprising since the regularization problems were part of the motivation for developing the method. However, there are some issues that must be taken into account when implementing LSTRS to treat ill-posed problems. As we saw in Section 2, for these problems, the smallest eigenvalues of H are clustered and close to zero and because of the interlacing property the smallest eigenvalues of B_α will also be clustered and small for certain

Figure 4: Viscoacoustic wave equation, $n = 500$.

values of Δ . Computing small eigenvalues with a method that relies on matrix-vector products with the original matrix is likely to fail since the multiplication will annihilate components precisely in the direction of the eigenvectors of interest. For this reason a Tchebyshev preconditioner was used to map the small eigenvalues to more significant values. In this approach instead of the eigenvalues of a given matrix B , we compute the eigenvalues of $T_n(B)$ where $T_n(z)$ is a Tchebyshev polynomial of degree n . The eigenvectors of B are then recover via Rayleigh quotient. Another issue for ill-posed problems concerns the computation of clustered eigenvalues which poses difficulties for any Lanczos-type method. We have not addressed this problem yet.

The occurrence of an interior solution when $A^T A$ is positive definite in regularization problems deserves a special comment. In this case the solution of the trust-region subproblem corresponds to the least squares solution of the original problem. This solution is contaminated by noise and is of no interest. When we detect an interior solution we have taken the simple approach of reducing the trust-region radius and restart the method. It is worth noticing that if we knew that the noise level in the data is low, then if λ is close to zero when we detect an interior solution, we could approximate the least squares solution by x satisfying (7) since this would be a reasonable approximation to $x = -H^{-1}g$. Note that in this case it would not be necessary to solve a linear system to obtain the solution.

5 Numerical Results

In this section we present numerical experiments to illustrate the performance of LSTRS on regularization problems from different sources, including both test problems and real applications. We used a Matlab version of LSTRS running under MATLAB 5.3 using Mexfile interfaces to access ARPACK [16] and also the routines to compute matrix–vector products in some of the examples. We ran our experiments on a SUN Ultrasparc 2 with a 200 MHz processor and 256 Megabytes of RAM running Solaris 5.6. The floating point arithmetic was IEEE standard double precision with machine precision $2^{-52} \approx 2.2204 \cdot 10^{-16}$.

We present three sets of experiments. In §5.1 we describe the results obtained on test problems from the Regularization Tools package [10]. In §5.2 we present an inverse interpolation problem on real data. In §5.3 we present a model seismic inversion problem on a standard data set.

5.1 Problems from the Regularization Tools package

In this section we will present the results of LSTRS on problems from the Regularization Tools package [10]. This package consists of a set of Matlab routines for the analysis of discrete ill–posed problems along with test problems that are easy to generate. All the test problems come from the discretization of a Fredholm integral equation of the first kind

$$\int_a^b K(s, t)f(t)dt = g(s),$$

where $K(s, t)$ is the kernel. The problem is to compute the unknown function $f(t)$ given $g(s)$ and $K(s, t)$.

In all cases we solve a quadratically constrained least squares problem (3) where A comes from the discretization of $K(s, t)$ and $b = g(s_i)$ at discrete points $s_i \in [a, b]$. In most cases, Ax was *different* from b . Unless otherwise specified we used $\Delta = \|x_{IP}\|$ as trust–region radius. When available, x_{IP} is the exact solution $f(t)$ evaluated at discrete points $t_i \in [a, b]$. In ARPACK, we used nine Lanczos basis vectors with seven shifts on each implicit restart. The required accuracy for the eigenpairs was 10^{-2} . The initial vector for the Lanczos factorization was a randomly generated vector that remains fixed in all the experiments. We solved the trust–region subproblems to a relative accuracy of $\left| \frac{\|x\| - \Delta}{\Delta} \right| < 10^{-2}$. We also solved the problems to a higher accuracy but this was computationally more expensive and did not seem to improve the accuracy of the regularized solution x with respect to the exact solution x_{IP} . In Table 1, we present the results for a subset of problems from [10].

Several observations are in order about Table 1. In the first place we observe on the third and fourth column that in all cases, LSTRS solved the trust–region

Problem	Dim.	Δ	$\ x\ $	$\frac{\ x-x_{IP}\ }{\ x_{IP}\ }$	MV Prods.	Iter.
Ill heat	300	4.2631	4.2527	3.5684e-01	1721	8
Ill heat	1000	7.7829	7.7497	2.6900e-01	967	8
Well heat	300	4.2631	4.2958	9.1853e-02	1049	4
ilaplace	195	2.7629	2.7362	1.8537e-01	349	4
parallax	300	5.0000	5.0421	–	869	10
phillips	300	2.9999	2.9869	2.6883e-02	521	6
phillips	1000	3.0000	2.9839	3.3607e-02	575	6
shaw	300	17.2893	17.2467	6.0625e-02	510	6
shaw	1000	31.5659	31.6002	5.2847e-02	423	5

Table 1: Results of LSTRS on test problems from the Regularization Tools package.

subproblem to the prescribed accuracy. The quality of the regularized solution or a measure of how well this solution approximates the exact solution x_{IP} is reported in the fifth column. We see that in most cases we obtain a reasonable agreement with relative errors of order 10^{-2} . For problems ill-conditioned **heat** (inverse heat equation) and **ilaplace** (inverse Laplace transformation), the error is higher and since in these cases we have a highly oscillatory solution, this suggests that we should solve the trust-region subproblem with a constraint of the form $\|Lx\|$ for a differential operator L . For these problems, our interest was to test LSTRS and not the regularization approach, and therefore we did not analyze each problem separately. We also observe that although the number of matrix-vector products is not too low in comparison with the dimension of the problem, this number does not seem to increase significantly with the dimension and in some cases it actually decreases.

5.2 An Inverse Interpolation Problem

The 2-D linear interpolation problem consists in using a linear interpolant to find the values of a function at arbitrary points given the values of the function at equally spaced points. If $A \in \mathbb{R}^{m \times n}$ represents the interpolant and $c \in \mathbb{R}^n$ contains the irregular spaced points then the values of the function at c are given by $z = Ac$.

A more interesting problem is the inverse interpolation problem: finding the values of the function on a regular grid of points from which we can extract given values of the function at irregularly spaced points by linear interpolation. We can pose the 2-D inverse interpolation problem as a least squares problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\| \quad (8)$$

where $A \in \mathbb{R}^{m \times n}$ is the 2-D linear interpolant and $b \in \mathbb{R}^m$ contains the function values at irregular spaced points.

To illustrate the performance of LSTRS on this kind of problem we will use the example of constructing a depth map of the Sea of Galilee, a fresh water lake in Israel. The data consists of triplets x_i, y_i, z_i $i = 1, \dots, 132044$ representing coordinates on the plane and depth, respectively. The data was collected from a ship using an echo sounder. There is noise in the data coming from different sources, including malfunctioning equipment that reported zero depths at points in the middle of the lake and the fact that the measurements were taken at different times of the year and therefore varied greatly from rainy season to dry season. A complete description of the data acquisition process can be found in [1]. In Figure 5 we show a view from above of a 3-D plot of the original data. The straight lines we observe in the figure are the measuring tracks, i.e. the data acquisition process was an additional source of noise. As Cl erbout points out in [5], an image of the sea should not include those lines.

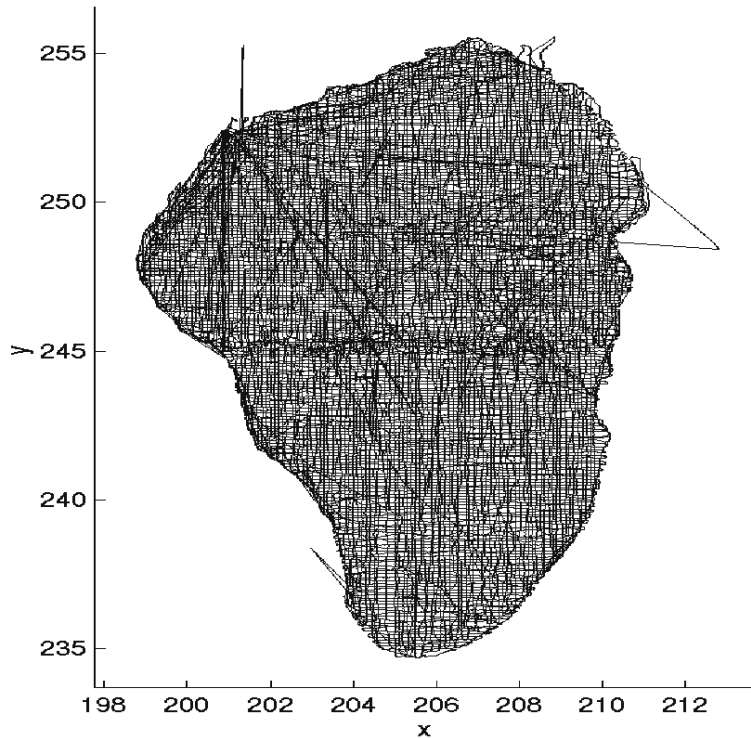


Figure 5: Sea of Galilee from the original data.

In our experiments the size of the grid was $n = 201 \times 201 = 40401$ and this is the number of unknowns when the 2-D grid is represented as a one-dimensional vector. The number of rows in A was $m = 132044$. This matrix was ill-conditioned and was not available explicitly, but we could compute the action of A and A^T on vectors by

means of FORTRAN routines. In all the experiments, we set $\varepsilon_\Delta = 10^{-3}$. The size of the Lanczos basis was five and we applied three shifts on each implicit restart. Therefore, the storage requirement was essentially five vectors of length 40401.

We posed the trust–region subproblem as

$$\begin{aligned} \min \quad & \frac{1}{2}x^T A^T A x - (A^T b)^T x \\ \text{s.t.} \quad & \|Lx\| \leq \Delta \end{aligned} \quad (9)$$

where L is either the identity matrix or a Helmholtz operator. We ran several experiments for different trust–region radii since in this application we did not have *a priori* information about the size or smoothness of the desired solution. Figure 6 shows the result for $\Delta = 6000$ and $L = I$, which we call a standard trust–region subproblem. We see that this image still shows the measurement tracks and does not reveal any known features of the depth distribution of the lake. The contour plot (not shown) is very rough and presents many oscillations. This suggested the need to introduce a constraint on the smoothness of the solution.

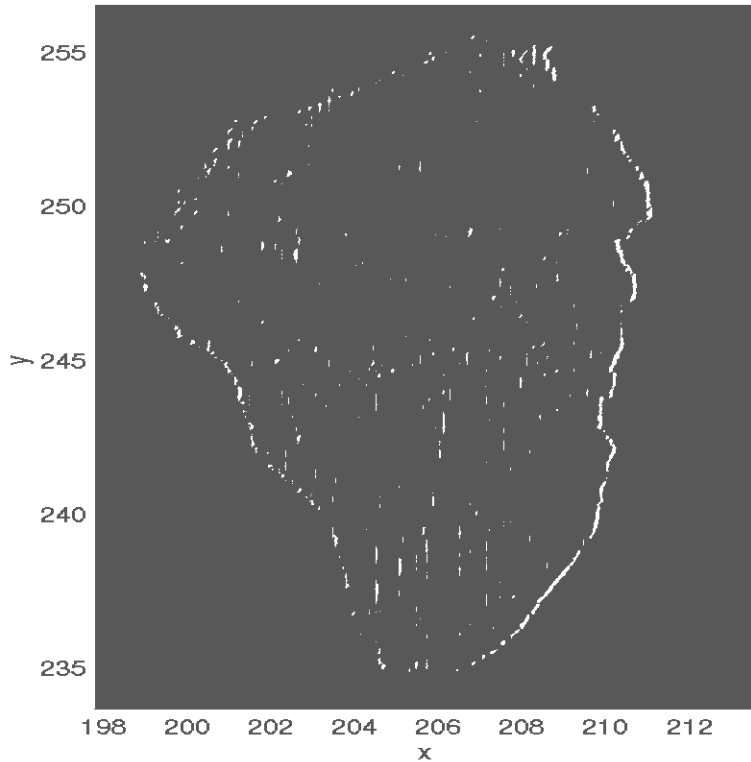


Figure 6: Sea of Galilee, regularizing with standard trust–region subproblem.

We then solved the trust–region subproblem with $L = M^p$ a Helmholtz operator, and $\Delta = 26000$. Figure 7 shows the solution for $p = 0.3$. In this image we were

able to identify some of the features reported in [1], such as the shallow areas in the Northeast, a scarp in the Southeast and a more prominent scarp in the Southwest.

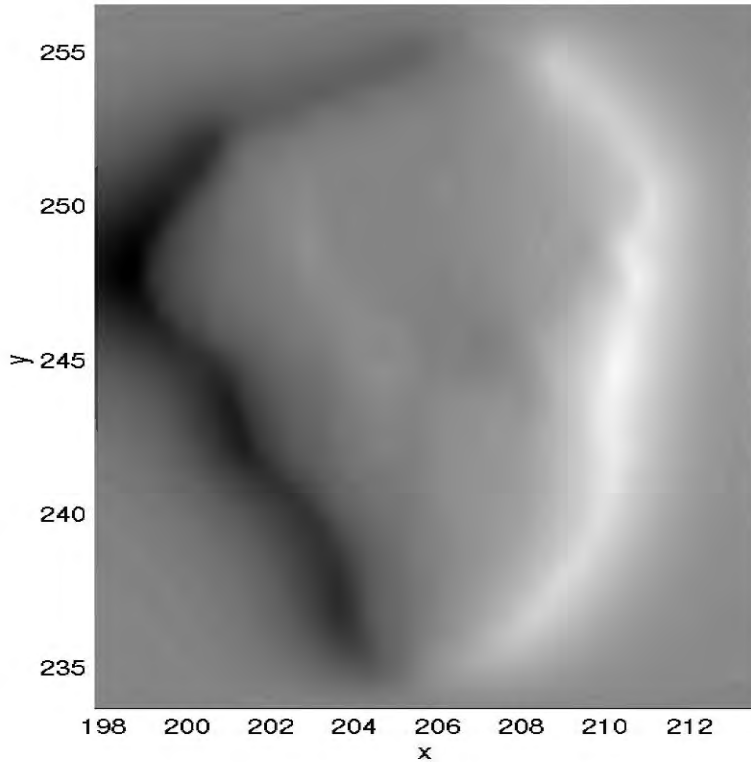


Figure 7: Sea of Galilee, regularizing with constraint on smoothness.

We also tried the approach of solving a standard trust–region subproblem and applying the Helmholtz operator *a posteriori*. We call this approach *post smoothing*. We tried the post smoothing approach with $p = -1$, for $\Delta = 23423$ which is the norm of x when $\|Lx\| = 26000$ and we obtained an image very similar to the one in Figure 7. We also used this approach for $\Delta = 6000$, obtaining the image in Figure 8 which clearly shows the features mentioned before. Table 2 shows that the post smoothing approach is less expensive than using a constraint on the smoothness. There are two reasons for this difference in efficiency. The first one is that the matrix–vector products when $L = M^p$ are more expensive than the matrix–vector products when $L = I$. The second reason is that the smallest eigenvalues of B_α are close to the smallest eigenvalues of $L^{-T}A^TAL^{-1}$ and these are more clustered for $L = M^p$ than for $L = I$. This causes slow convergence of the IRLM. We note however, that the cost is not too high in either approach relative to the dimension of the problem.

The post smoothing approach has the drawback that we do not know its physi-

cal meaning, and a closer look at the result shows that the post smoothing is causing a high degree of perturbation on the depths since the lowest point is known to be at around 256m below sea level and the lowest point in Figure 8 is -45m. Another interesting aspect of this particular regularization approach is that it produces very smooth solutions, which was also noted in [12], and this causes for example the resulting regular grid to miss the true deepest point located approximately at coordinates (207, 247), and yields a deepest point located at approximately (205, 245) for Figure 8. It is quite remarkable though how the known features of the lake are clearly present in this image.

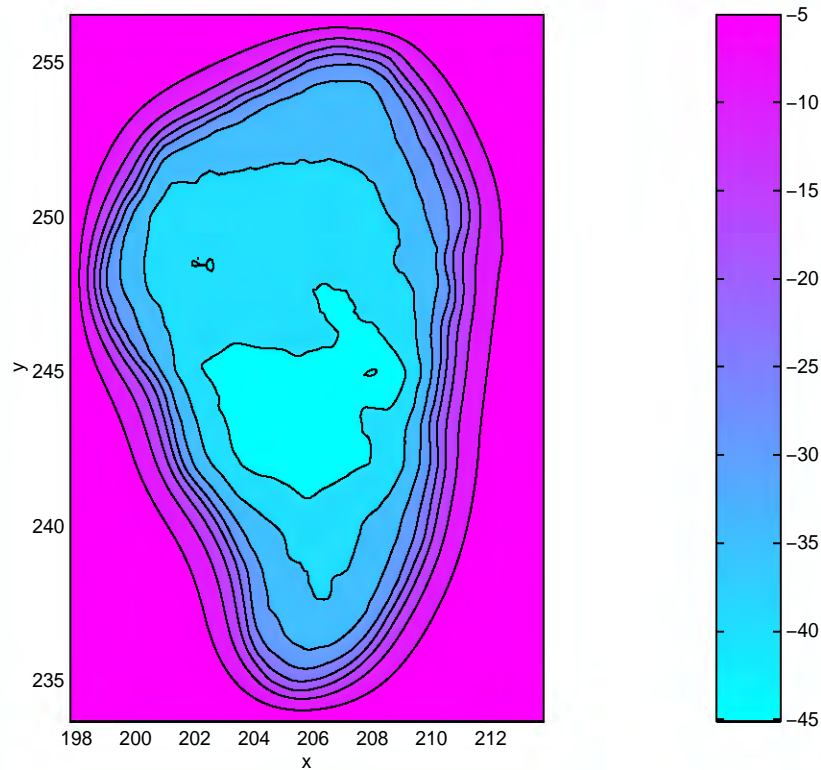


Figure 8: Sea of Galilee. Regularizing with standard constraint and post smoothing.

5.3 A Model Seismic Inversion Problem

We also solved the quadratically constrained least squares problem (3) where the problem comes from the discretization of the linear viscoacoustic model. As explained in [2], this model describes the behavior of an anelastic fluid, in which the strain response to a change of stress is linear but not completely instantaneous. A relaxation function $G(t, \vec{x})$ is used to express the stress-strain relation. The equa-

Dimension: 40401 Storage: 5 vectors	Δ	$\ x_*\ $	LSTRS Iter.	MV Prods.	CPU time (min.)
TRS with post smoothing	23423	23423	4	206	1.40
Constraint on smoothness	26000	25980.61	15	723	22.05

Table 2: Performance of LSTRS on an Inverse Interpolation Problem.

tions of motion relate $G(t, \vec{x})$, the material density $\rho(\vec{x})$, the pressure (stress) $p(t, \vec{x})$, the particle velocity, and the body force source $f(t, \vec{x})$ in the following way:

$$\begin{aligned}
 p_{,t}(t, \vec{x}) &= -\dot{G}(t, \vec{x}) * \nabla \cdot \vec{v}(t, \vec{x}) + f(t, \vec{x}) \\
 \vec{v}_{,t}(t, \vec{x}) &= -\frac{1}{\rho(\vec{x})} \nabla p(t, \vec{x})
 \end{aligned}$$

where $p = 0$, $\vec{x} = 0$ for $t \ll 0$.

This equation is used in [2] to model the propagation of waves in aquifer media using the relaxation function for a standard linear fluid. The experiment was performed using the AVO data set [14], a standard data set for testing inversion methods. The dimensions of the problem are $m = n = 121121$. Table 3 shows the result obtained when we used LSTRS to solve the trust-region subproblem to an accuracy of $\left| \frac{\|x\| - \Delta}{\Delta} \right| < 10^{-5}$ using four Lanczos basis vectors. The method is very efficient for small Δ since in this case the smallest eigenvalue of B_α is well separated from the rest and the IRLM is very efficient for computing it. For larger Δ , the eigenvalues of interest become more clustered and it takes more iterations for the IRLM to converge.

Dimension	Δ	$\ x\ $	LSTRS Iter.	MV Prods.	Storage
121121	0.5	0.5	2	15	4 vectors

Table 3: Performance of LSTRS on the viscoacoustic wave equation.

In Table 3 we can observe the low storage and low number of matrix-vectors products required to solve the problem to a very high accuracy.

6 Conclusions

We considered the problem of regularizing large-scale discrete forms of ill-posed problems arising in several applications. We posed the regularization problem as a quadratically constrained least squares problem and showed the relationship of this approach to Tikhonov regularization and to the trust-region subproblem and analyzed the latter in the ill-posed case.

We have presented numerical results obtained when we used the recently developed method LSTRS for the large-scale trust-region subproblem to solve regularization problems from a wide variety of applications including problems with real data. The method requires solving a sequence of large-scale eigenvalue problems which is accomplished with a variant of the Lanczos method.

LSTRS is particularly suitable for large-scale discrete ill-posed problems for which it computed regularized solutions close to the desired exact solution using limited storage and moderate computational effort in general. For real applications the method used both less storage than the Conjugate Gradient method and low number of matrix-vector products, while providing systematic stopping criteria. Therefore, the method showed to be competitive with other techniques for solving these problems in particular when the desired solution has small norm. Large norms give rise to clustered eigenvalues and this poses difficulties for the Lanczos method on which LSTRS relies. An important feature of LSTRS is that it computes both the solution and the Tikhonov regularization parameter from the prescribed norm.

Although further improvement is needed, such as the used of preconditioning for the eigenvalue problems, LSTRS proved to be a promising method for the numerical treatment of large-scale ill-posed problems in which the norm of the desired solution is prescribed.

Acknowledgements. We would like to thank Chao Yang for making his Matlab code for the Tchebyshev preconditioner available to us. We would also like to express our deep gratitude to Bill Symes who provided us with the data, codes for the 2-D linear interpolant and interpretation of the results in Section 5.2. Bill also provided the Model Seismic Problem in Section 5.3. We also thank Professor Zvi Ben-Avraham from Tel-Aviv University, for making the data on the Sea of Galilee available and for providing important references.

References

- [1] Z. BEN-AVRAHAM, G. AMIT, A. GOLAN and Z.B. BEGIN. The bathymetry of Lake Kinneret and its structural significance, *Isr. J. Earth Sci.*, 39:77–84, 1992.

- [2] J.O. BLANCH, W.W. SYMES and R.J. VERSTEEG. A numerical study of linear viscoacoustic inversion. In R.G. Keys and D.J. Foster, editors. *Comparison of seismic inversion methods on a single real data set*, pages 13–44. Society of Exploration Geophysicists, Tulsa, 1998.
- [3] Å. BJÖRCK, E. GRIMME and P. VAN DOOREN. An implicit shift bidiagonalization algorithm for ill-posed systems, *BIT*, 34:510–534, 1994.
- [4] D. CALVETTI, L. REICHEL and Q. ZHANG. Iterative exponential filtering for large discrete ill-posed problems, To appear in *Numer. Math.*
- [5] J. CLÆRBOUT. GEOPHYSICAL ESTIMATION BY EXAMPLE: Environmental soundings image enhancement: Multidimensional autoregression. Available from <http://sepwww.stanford.edu/sep/jon/index.html>.
- [6] G.H. GOLUB and C.F. VAN LOAN. *Matrix Computations*, Johns Hopkins University Press, Baltimore, third ed., 1996.
- [7] G.H. GOLUB and U. VON MATT. Quadratically constrained least squares and quadratic problems, *Numer. Math.*, 59:561–580, 1991.
- [8] M. HANKE and P.C. HANSEN. Regularization methods for large-scale problems, *Surveys of Mathematics for Industry*, 3:253–315, 1993.
- [9] P.C. HANSEN. The Discrete Picard Condition for discrete ill-posed problems, *BIT*, 30:658–672, 1990.
- [10] P.C. HANSEN. Regularization Tools: a MATLAB package for analysis and solution of discrete ill-posed problems, *Numer. Algo.*, 6:1–35, 1994.
- [11] P.C. HANSEN. Rank-Deficient and Discrete Ill-Posed Problems. Doctoral Dissertation, UNI•C, Technical University of Denmark, Lyngby, Denmark, 1996.
- [12] P.C. HANSEN, M. JACOBSEN, J.M. RASMUSSEN. and H. SØRENSEN. The PP-TSVD algorithm for image restoration problems. In P.C. Hansen, B.H. Jacobsen and K. Mosegaard, editors. *Inverse Methods II*, Springer, to appear. May 1999.
- [13] L. KAUFMAN and A. NEUMAIER. Image reconstruction through regularization by envelope guided Conjugate Gradients, AT&T Bell Laboratories, 4–14, 1994
- [14] R.G. KEYS and D.J. FOSTER. A data set for evaluating and comparing seismic inversion methods. In R.G. Keys and D.J. Foster, editors. *Comparison of seismic inversion methods on a single real data set*, pages 1–12. Society of Exploration Geophysicists, Tulsa, 1998.

- [15] M.E. KILMER and D.P. O'LEARY. Pivoted Cauchy-like preconditioners for regularized solution of ill-posed problems, *SIAM J. Sci. Comput.*, 21(1):88–110, 1999.
- [16] R.B. LEHOUCQ, D.C. SORESENSEN and C. YANG. *ARPACK User's Guide: Solution of Large Scale Eigenvalue Problems by Implicitly Restarted Arnoldi Methods*, SIAM, Philadelphia, 1998.
- [17] J.J. MORÉ and D.C. SORESENSEN. Computing a trust region step, *SIAM J. Sci. Stat. Comput.*, 4(3):553–572, 1983.
- [18] J. NAGY, V. PAUCA, R. PLEMMONS and T. TØRGERSEN. Space-varying restoration of optical images, Department of Mathematics and Computer Science, Wake Forest University, Winston-Salem, North Carolina, March 1996.
- [19] A. NEUMAIER. Solving ill-conditioned and singular linear systems: a tutorial on regularization, *SIAM Review*, 40(3):636–666, 1998.
- [20] B.N. PARLETT. *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs NY, 1980.
- [21] M. ROJAS. A Large-Scale Trust-Region Approach to the Regularization of Discrete Ill-Posed Problems. Ph.D. Thesis, Technical Report TR98-19, Department of Computational and Applied Mathematics, Rice University, Houston, May 1998.
- [22] M. ROJAS, S.A. SANTOS and D.C. SORESENSEN. A new matrix-free algorithm for the large-scale trust-region subproblem. Technical Report TR99-19, Department of Computational and Applied Mathematics, Rice University, Houston, September 1999. Submitted.
- [23] D.C. SORESENSEN. Newton's method with a model trust region modification, *SIAM J. Numer. Anal.*, 19(2):409–426, 1982.
- [24] D.C. SORESENSEN. Implicit application of polynomial filters in a k-step Arnoldi method, *SIAM J. Matrix Anal. Appl.*, 13(1):357–385, 1992.
- [25] A.N. TIKHONOV. Regularization of incorrectly posed problems, *Soviet Math.*, 4:1624–1627, 1963.