Regularization Algorithms Based on Total Least Squares

Per Christian Hansen* Dianne P. O'Leary†

Abstract

Discretizations of inverse problems lead to systems of linear equations with a highly ill-conditioned coefficient matrix, and in order to compute stable solutions to these systems it is necessary to apply regularization methods. Classical regularization methods, such as Tikhonov's method or truncated *SVD*, are not designed for problems in which both the coefficient matrix and the right-hand side are known only approximately. For this reason, we develop *TLS*-based regularization methods that take this situation into account.

Here, we survey two different approaches to incorporation of regularization, or stabilization, into the *TLS* setting. The two methods are similar in spirit to Tikhonov regularization and truncated *SVD*, respectively. We analyze the regularizing properties of the methods and demonstrate by numerical examples that in certain cases with large perturbations, these new methods are able to yield more accurate regularized solutions than those produced by the standard methods.

1 Discrete Ill-Posed Problems

In this paper we study linear, and possibly overdetermined, systems of equations $Ax \approx b$ whose $m \times n$ coefficient matrix A (with $m \geq n$) is very ill conditioned. We restrict our attention to the important case where all the singular values of A decay gradually to zero, i.e., with no particular gap in the spectrum.

Such ill-conditioned linear systems arise frequently in connection with discretizations of ill-posed problems, such as Fredholm integral equations of the first kind, and the term *discrete ill-posed problem* is sometimes used to characterize these systems. For more details about the underlying theory see, e.g., [2], [3], [6], [8] and the references therein. Suffice it here to say that the gradual decay of the singular values of A is an intrinsic property of discretizations of many ill-posed problems.

For discrete ill-posed problems, the ordinary least squares solution x_{LS}, as well as the ordinary total least squares solution x_{TLS}, are hopelessly contaminated by noise in the directions corresponding to the small singular values of A or $(A\,,\,b)$. Because of this, it is necessary to compute a *regularized solution* in which the effects of the noise are filtered out. Surveys of regularization methods for discrete ill-posed problems are given in [6] and [8].

The filtering is often done either by truncation of the small singular values of A or by Tikhonov's method.

If $A = \sum_{i=1}^{n} u_i'\, \sigma_i'\, v_i'^T$ is the *SVD* of A, then the truncated *SVD* (*TSVD*) solution x_k,

*Department of Mathematical Modelling, Building 305, Technical University of Denmark, DK-2800 Lyngby, Denmark (pch@mm.dtu.dk)

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA (oleary@cs.umd.edu). This work was supported by the U.S.A. National Science Foundation under Grant CCR 95-03126.

with truncation parameter k, is given by

$$x_k = \sum_{i=1}^{k} \frac{u_i'^T b}{\sigma_i'} v_i', \quad k \le n. \tag{1}$$

In the Tikhonov method, a side constraint $\|Lx\|_2 \le \delta$ is added to the least squares formulation:

$$\min \|Ax - b\|_2 \quad \text{s.t.} \quad \|Lx\|_2 \le \delta. \tag{2}$$

If $L = I_n$ then the problem is in *standard form*, but it is often advantageous to choose L as a discrete approximation to a derivative operator. A solution to this optimization problem solves the system of equations

$$(A^T A + \lambda L^T L)\, x = A^T b \tag{3}$$

with $\lambda \ge 0$. It can be shown that the solution x_δ to (2) is identical to the solution to (3) for an appropriately chosen λ, and there is a monotonic relation between the parameters δ and λ. In the standard-form case ($L = I_n$), the Tikhonov solution is given by

$$x_\delta = \sum_{i=1}^{n} \frac{\sigma_i'^2}{\sigma_i'^2 + \lambda} \frac{u_i'^T b}{\sigma_i'} v_i'. \tag{4}$$

For both the truncated *SVD* and the Tikhonov algorithm, the regularized solution is a filtered version of the ordinary least squares solution. The *filter factors* (the ratio between the coefficients of v_i' in the computed solution and the exact solution x_{LS}) are zeros and ones for the *TSVD* solution and $\sigma_i'^2/(\sigma_i'^2 + \lambda)$, for the Tikhonov solution.

Most regularization methods used today assume that the errors are confined to the right-hand side b. Although this is true in many applications there are problems in which also the coefficient matrix A is not precisely known. For example, A may be available only by measurement, or may be an idealized approximation of the true operator. Discretization typically also adds some errors to the matrix A. Hence, there is a need for regularization methods that take into account the errors in A and their size relative to those in b.

In this paper we survey two such regularization methods in the *TLS* setting. One is analogous to the Tikhonov-regularized solution and the other to the truncated SVD, but both allow errors in the entries of A. These methods have been developed recently in [4] and [5]. We discuss the regularizing effects of these methods and illustrate by numerical examples that they can be superior to the classical regularization methods.

2 Regularized *TLS*

The first *TLS*-based regularization method is based on the Tikhonov formulation (2). In the *TLS* setting, we add the bound $\|Lx\|_2 \le \delta$ to the ordinary *TLS* problem, and the *regularized TLS* (*R-TLS*) problem thus becomes

$$\min_{A_0, x} \|(A\,,\, b) - (A_0\,,\, b_0)\|_F \quad \text{s.t.} \quad b_0 = A_0 x, \quad \|Lx\|_2 \le \delta. \tag{5}$$

The corresponding Lagrange formulation is

$$\mathcal{L}(A_0, x, \mu) = \|(A\,,\, b) - (A_0\,,\, A_0 x)\|_F^2 + \mu\,(\|x\|_2^2 - \delta^2)\,, \tag{6}$$

where μ is the Lagrange parameter. The *R-TLS* solution $\bar{x}_\delta$ to (5) is characterized by the following theorem from [5].

Theorem 2.1. *The regularized TLS solution to (5) is a solution to the problem*

$$(A^T A + \lambda_I I_n + \lambda_L L^T L)\, x = A^T b \tag{7}$$

where the parameters λ_I and λ_L are given by

$$\lambda_I = -\frac{\|b - Ax\|_2^2}{1 + \|x\|_2^2} \tag{8}$$

$$\lambda_L = \mu\,(1 + \|x\|_2^2) \tag{9}$$

and where μ is the Lagrange multiplier in (6). Moreover, the TLS residual satisfies

$$\|(A\,,\, b) - (A_0\,,\, b_0)\|_F^2 = -\lambda_I. \tag{10}$$

2.1 The Standard-Form Case

In the standard-form case ($L = I_n$), Eq. (7) simplifies to

$$(A^T A + \lambda_{IL} I_n)\, x = A^T b \tag{11}$$

with $\lambda_{IL} = \lambda_I + \lambda_L$. In this case, the standard-form *R-TLS* solution $\bar{x}_\delta$ and the standard-form Tikhonov solution x_δ have a close relationship which is proved in [5].

Theorem 2.2. *For a given value of δ, the solutions $\bar{x}_\delta$ and x_δ are related as follows, where σ_{n+1} denotes the smallest singular value of $(A\,,\, b)$:*

δ	solutions	λ_{IL}
$\delta < \|x_{\mathrm{LS}}\|_2$	$\bar{x}_\delta = x_\delta$	$\lambda_{IL} > 0$
$\delta = \|x_{\mathrm{LS}}\|_2$	$\bar{x}_\delta = x_\delta = x_{\mathrm{LS}}$	$\lambda_{IL} = 0$
$\|x_{\mathrm{LS}}\|_2 < \delta < \|x_{\mathrm{TLS}}\|_2$	$\bar{x}_\delta \neq x_\delta = x_{\mathrm{LS}}$	$-\sigma_{n+1}^2 < \lambda_{IL} < 0$
$\delta \geq \|x_{\mathrm{TLS}}\|_2$	$\bar{x}_\delta = x_{\mathrm{TLS}},\ x_\delta = x_{\mathrm{LS}}$	$\lambda_{IL} = -\sigma_{n+1}^2$

We conclude that as long as $\delta \leq \|x_{\mathrm{LS}}\|_2$, which is normally the case in regularization problems where $\|x_{\mathrm{LS}}\|_2$ is very large, then regularized *TLS* produces solutions that are identical to the Tikhonov solutions. In other words, replacing the *LS* residual with the *TLS* residual in the Tikhonov formulation has no effect when $L = I_n$ and $\delta \leq \|x_{\mathrm{LS}}\|_2$.

We remark that since $\|x_{\mathrm{TLS}}\|_2 \geq \|x_{\mathrm{LS}}\|_2$ (see [12, Corollary 6.2]) there is usually a nontrivial set of "large" δ's for which the multiplier λ_{IL} is negative. The corresponding *R-TLS* solutions $\bar{x}_\delta$ can be expected to be even more dominated by errors than the least squares solution x_{LS}.

2.2 The General-Form Case

In many applications, it is necessary to choose a matrix L different from the identity matrix, and often L is chosen to represent the first or second derivative operator. In this case, the *R-TLS* solution $\bar{x}_\delta$ is different from the Tikhonov solution whenever the residual $b - Ax$ is different from zero, since both λ_I and λ_L are nonzero.

Notice that λ_L is always positive, as long as $\delta < \|x_{\mathrm{TLS}}\|_2$ (because the Lagrange parameter μ is positive for these values of λ). On the other hand, λ_I is always negative, and thus adding some *de-regularization* to the solution. Statistical aspects of a negative regularization parameter in Tikhonov's method are discussed in [9].

For a given δ, there are usually several pairs of parameters λ_I and λ_L, and thus several solutions x, that satisfy relations (7)–(9), but only one of these satisfies the optimization problem (5). According to (10), this is the solution that corresponds to the smallest value of $|\lambda_I|$. The following relations hold:

δ	solutions	λ_I	λ_L
$\delta < \|Lx_{\mathrm{TLS}}\|_2$	$\bar{x}_\delta$	$\lambda_I < 0 \quad \partial\lambda_I/\partial\delta > 0$	$\lambda_L > 0 \quad \partial\lambda_L/\partial\delta < 0$
$\delta \geq \|Lx_{\mathrm{TLS}}\|_2$	$\bar{x}_\delta = x_{\mathrm{TLS}}$	$\lambda_I = -\sigma_{n+1}^2$	$\lambda_L = 0$

We note that if the matrix $\lambda_I I_n + \lambda_L L^T L$ is positive definite, then the *R-TLS* solution corresponds to a Tikhonov solution for which the seminorm $\|Lx\|_2$ in (3) is replaced with the Sobolev norm $(\lambda_I\|x\|_2^2 + \lambda_L\|Lx\|_2^2)^{1/2}$. If $\lambda_I I_n + \lambda_L L^T L$ is indefinite or negative definite then there is no equivalent interpretation.

2.3 Computational Aspects

To compute the *R-TLS* solutions for $L \neq I_n$, we have found it most convenient to avoid explicit use of δ; instead we use λ_L as the free parameter, fixing its value and then computing the value of λ_I that satisfies (8) and is smallest in absolute value. As shown in [5], the corresponding value of δ can then easily be computed from the relation

$$\lambda_L\,\delta^2 = b^T(b - Ax) + \lambda_I. \tag{12}$$

We now discuss how to solve (7) efficiently for many values of λ_I and λ_L. We assume that the matrix L is a banded matrix, which is often the case when L approximates a derivative operator. The key to efficiency is then to reduce A to bidiagonal form B by means of orthogonal transformations: $H^T A\,K = B$. The orthogonal right-transformations should also be applied to L, and simultaneously we should apply orthogonal transformations to L from the left in order to maintain its banded form. It is convenient to use sequences of Givens transformations to form J, H and K, since this gives us the most freedom to retain the banded form of $C = J^T L\,K$.

Once B and C have been computed, we note that (7) is equivalent to the following least squares problem

$$\min \left\| \begin{pmatrix} B \\ \lambda_L C \\ \hat{\imath}\,\lambda_I I_n \end{pmatrix} (K^T x) - \begin{pmatrix} H^T b \\ 0 \\ 0 \end{pmatrix} \right\|_2 \tag{13}$$

where $\hat{\imath}$ is the imaginary unit. Since λ_I changes more frequently than λ_L in our approach, the next step is to reduce the submatrix $\begin{pmatrix} B \\ \lambda_L C \end{pmatrix}$ to bidiagonal form $\hat{B}$ by means of Givens rotations, along the same lines as in Elden's algorithm [1, Section 5.3.4]. This changes $\begin{pmatrix} H^T b \\ 0 \end{pmatrix}$ into $\hat{d}$, and thus we arrive at the problem

$$\min \left\| \begin{pmatrix} \hat{B} \\ \hat{\imath}\,\lambda_I I_n \end{pmatrix} (K^T x) - \begin{pmatrix} \hat{d} \\ 0 \end{pmatrix} \right\|_2 . \tag{14}$$

We are currently investigating stable and efficient numerical algorithms for solving (14).

3 Truncated *TLS*

The second *TLS*-based approach to regularization is inspired by the *TSVD* method in which the small singular values of A are discarded. In the truncated *TLS* (*T-TLS*) method the key idea is to neglect the small singular values of $(A\,,\,b)$, by setting those below a given threshold to zero; see, for example, [12]. The idea to apply the *T-TLS* method to discrete ill-posed problems where there is no particular gap in the singular value spectrum, and where some of the larger singular values may also be discarded, was proposed in [4].

The details of the *T-TLS* method are as follows. Let the *SVD* of $(A\,,\,b)$ be given by

$$(A\,,\,b) = U\,\Sigma\,V^T = \sum_{i=1}^{n+1} u_i\,\sigma_i\,v_i^T \tag{15}$$

and assume that σ_k is the smallest nonzero singular value that we wish to retain in the *T-TLS* solution. As is usual in *TLS* problems, we also assume that σ_k is separated from σ_{k+1}. If we partition the $(n+1)\times(n+1)$ matrix V in (15) such that

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}, \quad V_{11} \in \Re^{n\times k}$$

then the *T-TLS* solution is given by

$$\bar{x}_k = -V_{12}V_{22}^{\dagger} = -V_{12}V_{22}^T\|V_{22}\|_2^{-2}. \tag{16}$$

Here $V_{22}^{\dagger} = V_{22}^T\|V_{22}\|_2^{-2}$ is the pseudoinverse of the $1\times(n-k+1)$ submatrix V_{22}.

The following relations follow immediately from (16):

$$\|\bar{x}_k\|_2 \;=\; \sqrt{\|V_{22}\|_2^{-2} - 1}\,, \tag{17}$$

$$\|(A\,,\,b) - (A_0\,,\,b_0)\|_F \;=\; \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_{n+1}^2}\,, \tag{18}$$

showing that the solution norm $\|\bar{x}_k\|_2$ increases monotonically with k while the *TLS* residual norm $\|(A\,,\,b) - (A_0\,,\,b_0)\|_F$ decreases monotonically with k.

3.1 The *T-TLS* Filter Factors (Standard Form)

One of the key results in [4] is an expression for the *T-TLS* solution $\bar{x}_k$ in terms of the *SVD* of A, rather than the *SVD* of $(A\,,\,b)$ as was previously the case. The advantage of using the *SVD* of A is that it immediately yields the *T-TLS* filter factors and thus lets us quantify the regularizing properties of the *T-TLS* solution. We emphasize that the following theorem, whose proof can be found in [4], holds generally, and not just for discrete ill-posed problems. Thus, it supplements Thm. 3.8 in [12].

Theorem 3.1. *Assume that the singular values of A and $(A\,,\,b)$ are simple. Then the T-TLS solution can be written as*

$$\bar{x}_k = \sum_{i=1}^{n} f_i\,\frac{u_i'^T b}{\sigma_i'}\,v_i'^T, \tag{19}$$

where the filter factors f_i corresponding to $u_i'^T b \neq 0$ and $\sigma_i' \neq 0$ are given by

$$f_i = \sum_{j=1}^{k} \frac{v_{n+1,j}}{\|V_{22}\|_2^2}\,\frac{\sigma_i'^2}{\sigma_j^2 - \sigma_i'^2}. \tag{20}$$

The filter factors for $i \le k$ corresponding to $u_i^T b \neq 0$ increase monotonically with i and satisfy

$$0 \le f_i - 1 \le \frac{\sigma_{k+1}^2}{\sigma_i'^2 - \sigma_{k+1}^2}\;. \tag{21}$$

The filter factors for $k < i \leq \operatorname{rank}(A)$ corresponding to $u_i^T b \neq 0$ satisfy

$$0 \leq f_i \leq \|V_{22}\|_2^{-2} \, \frac{\sigma_i'^2}{\sigma_k^2 - \sigma_i'^2} \; . \tag{22}$$

As an immediate consequence of the results in Thm. 3.1 we have $\|\bar{x}_k\|_2 \geq \|x_k\|_2$ for all k. Moreover, we obtain the following expression for the first k filter factors

$$1 \leq f_i \leq 1 + \frac{\sigma_{k+1}^2}{\sigma_i'^2} + \mathcal{O}\left(\frac{\sigma_{k+1}^4}{\sigma_i'^4}\right) \quad , \quad i = 1, \ldots, k \; ,$$

showing that the larger the ratio between σ_i' and σ_{k+1}, the closer the bound on f_i is to 1. Similarly, for the last $n-k$ filter factors we obtain

$$0 \leq f_i \leq \|\bar{V}_{22}\|_2^{-2} \frac{\sigma_i^2}{\bar{\sigma}_k^2} \left(1 + \mathcal{O}\left(\frac{\sigma_i^2}{\bar{\sigma}_k^2}\right)\right), \quad i = k+1, \ldots, n \; ,$$

showing that the smaller the ratio between σ_i' and σ_k, the closer f_i is to zero.

Hence, Thm. 3.1 guarantees that the first k filter factors will be close to one and that the last $n-k$ filter factors will be small, even in the case where there is no gap in the singular value spectrum, provided that $\|\bar{V}_{22}\|_2$ is not very small. As a consequence, the *T-TLS* solution is a regularized, or filtered, solution.

3.2 A Lanczos Bidiagonalization Algorithm

When the dimensions of A are not too large, one can compute the full *SVD* of $(A\,,\, b)$ and then experiment with various choices of k. This is particularly useful if no a priori estimate of a suitable k is known.

When the dimensions of A become large, this approach becomes prohibitive because the *SVD* algorithm is of complexity $\mathcal{O}(mn^2)$. We shall therefore describe an alternative technique that is much more suited for large-scale problems with $k \ll n$, which is indeed the case in most discrete ill-posed problems.

Our algorithm uses the Lanczos bidiagonalization process which computes approximations to the principal singular triplets of a matrix. Approximations to the ordinary *TLS* solution can thus be computed by applying the Lanczos bidiagonalization process to the matrix $(A\,,\, b)$, cf. [1, Section 7.6.5].

The *Lanczos T-TLS* algorithm proposed in [4] is based on Lanczos bidiagonalization of the matrix A rather than $(A\,,\, b)$. After k iterations, the Lanczos bidiagonalization process with starting vector $u_1 = b/\|b\|_2$ has produced two sets of vectors $U_k = (u_1, \ldots, u_{k+1})$ and $V_k = (v_1, \ldots, v_k)$ and a $(k+1) \times k$ bidiagonal matrix B_k such that

$$A\, V_k = U_k B_k \qquad \text{and} \qquad \beta_1 u_1 = b \; .$$

Thus, after k Lanczos iterations we can project the *TLS* problem onto the subspaces spanned by U_k and V_k, in the hope that for large enough k we have captured all the large singular values of A that are needed for computing a useful regularized solution. The projected *TLS* problem is equivalent to

$$\min \left\| U_k^T \Big((A\,,\, b) - (A_{0,k}\,,\, b_{0,k}) \Big) \begin{pmatrix} V_k & 0 \\ 0 & 1 \end{pmatrix} \right\|_F \quad \text{s.t.} \quad U_k^T A_{0,k}\, V_k\, y = U_k^T b_{0,k} \; ,$$

or

$$\min \|(B_k\,,\, \beta_1 e_1) - (B_{0,k}\,,\, e_{0,k})\|_F \qquad \text{subject to} \qquad B_{0,k}\, y = e_{0,k}\;, \tag{23}$$

where $e_1 = (1, 0, \ldots, 0)^T$, and $B_{0,k}$ and $e_{0,k}$ are generally full. Our algorithm reduces to the *LSQR* algorithm [10] if we require $B_{0,k} = B_k$ in each step.

In each Lanczos step we can now compute an approximate *T-TLS* solution by applying the ordinary *TLS* algorithm to the small-size problem in (23). For convenience, we can permute the vector $\beta_1 e_1$ in front of B_k such that, in each step, we merely need to compute the last singular triplet of the $(k+1) \times (k+1)$ upper bidiagonal matrix $(\beta_1 e_1\,,\, B_k)$. This can be done in $\mathcal{O}(k^2)$ operations by means of the *PSVD* algorithm [13].

We remark that it is easy to augment the above algorithm to include the computations of the *LSQR* algorithm [10]. Approximate *TSVD* solutions can thus be computed together with the approximate *T-TLS* solutions with little overhead.

3.3 The General-Form Case

In connection with the *T-TLS* and Lanczos *T-TLS* algorithms it may also be convenient to implicitly use regularization in general form with $L \neq I$. This is done in the same way as general-form regularization is treated numerically in other regularization methods. First transform the problem involving A, L and b into a standard-form problem. Then apply the *T-TLS* or Lanczos *T-TLS* algorithm to the standard-form problem to obtain a regularized solution. Finally, transform this solution back to the general-form setting. We omit the details here and refer to the discussion of the implementation details in [6, §4.3] and [8].

4 Numerical Results

In this section we present some numerical results that demonstrate the usefulness of the *R-TLS* and *T-TLS* methods. All computations were carried out in Matlab using the Regularization Tools package [7]. More elaborate tests can be found in [4] and [5], where results from tests with the Lanczos *T-TLS* algorithm can also be found.

It is a generally accepted fact that for small noise levels, we should not expect the ordinary *TLS* solution to differ much from the ordinary least squares solution; see [11]. The same observations are made in [4] and [5] for the *TLS*-based regularized solutions, and the numerical results presented below also support this observation. We emphasize that the precise meaning of "small" depends on the particular problem.

4.1 Illustration of the *R-TLS* Algorithm

The test problem we have chosen to illustrate the *R-TLS* algorithm is a discretization by means of Gauss-Laguerre quadrature of the inverse Laplace transform

$$\int_0^\infty \exp(-s\,t)\, f(t)\, dt = \frac{1}{2} - \frac{1}{s+1/2}, \qquad 0 \le s$$

$$f(t) = 1 - \exp(-t/2)$$

as implemented in the function `ilaplace(n,2)` in [7]. The matrix L approximates the first derivative operator. The dimensions are $m = n = 16$, the matrix A and the exact right-hand side $b^{\text{exact}} = A x^{\text{exact}}$ are scaled such that $\|A\|_F = \|b^{\text{exact}}\|_2 = 1$, and the perturbed right-hand side is generated as

$$b = (A + \sigma \|E\|_F^{-1} E) x^{\text{exact}} + \sigma \|e\|_2^{-1} e,$$

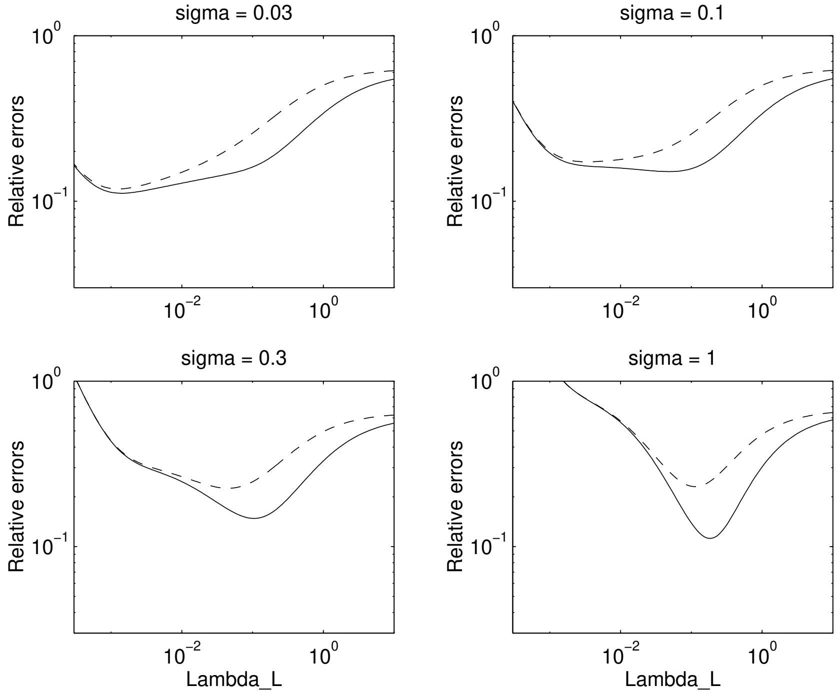

FIG. 1. *Plots of the relative errors in the Tikhonov solutions (dashed lines) and R-TLS solutions (solid lines) versus λ_L for four values of the noise level.*

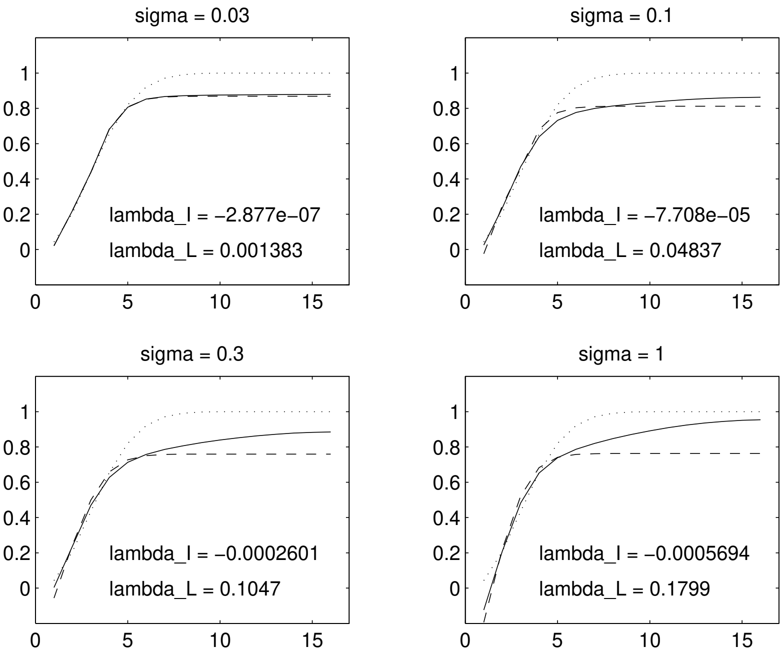

FIG. 2. *Plots of the "optimal" Tikhonov solutions (dashed lines) and R-TLS solutions (solid lines) for four values of the noise level. Also shown is the exact solution (dotted lines).*

where $x^{\rm exact}$ is the exact solution, and E and e are perturbations with unbiased normally distributed elements.

Figure 1 shows the relative errors $\|x^{\rm exact} - x_\delta\|_2/\|x^{\rm exact}\|_2$ and $\|x^{\rm exact} - \bar{x}_\delta\|_2/\|x^{\rm exact}\|_2$

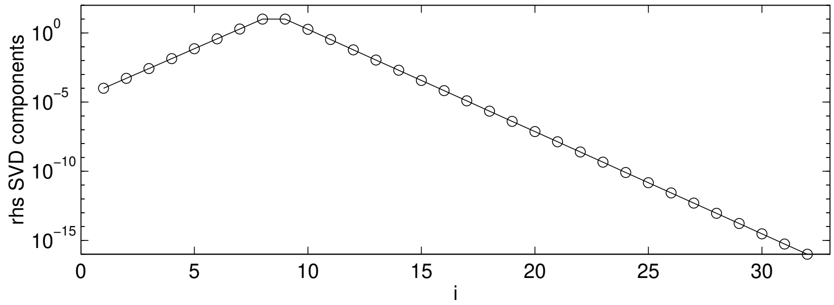

FIG. 3. *The SVD components $u_i'^T b^{\rm exact}$ of the exact right-hand side in the "artificially generated" test problem.*

in the Tikhonov and *R-TLS* solutions, respectively, for four values of the noise level:

$$\sigma = 0.03,\ 0.1,\ 0.3,\ 1.0.$$

We see that for small values of σ the two methods lead to almost the same minimum relative error, for almost the same value of λ_L. However, for larger value of σ, the minimum relative error for the *R-TLS* method is clearly smaller than that for Tikhonov's method, and it occurs for a larger value of λ_L. This shows the potential advantage of the *R-TLS* method, provided, of course, that a good estimate of the optimal regularization parameter can be found. This topic is outside the scope of the current paper.

In Fig. 2 we have plotted the "optimal" Tikhonov and *R-TLS* solutions, defined as the solutions that correspond to the minima of the curves in Fig. 1. In addition, we have plotted the exact solution $x^{\rm exact}$. Clearly, the addition of the term $\lambda_I I_n$ in (7) introduces a non-constant component in the right part of the plot of the regularized solution, and it is precisely this component that improves the *R-TLS* error, compared to the Tikhonov error.

4.2 Illustration of the *T-TLS* Algorithm

For the *T-TLS* method, it was found in [4] that the *T-TLS* solutions are superior to the other regularized solutions only when the *T-TLS* solution has large *SVD* components corresponding to the smallest retained singular values in the solution.

To illustrate this, we use an "artificially generated" test problem similar to the one used in [4]. The matrix A is 32×32 and comes from discretization of Phillip's test problem (cf. phillips in [7]) with kernel

$$K(s,t) = \begin{cases} 1 + \cos(\pi(s-t)/3)\ , & |s-t| < 3 \\ 0 & |s-t| \geq 3 \end{cases}$$

by means of the Galerkin method with ⎍ orthonormal basis functions. The exact solution $x^{\rm exact}$ is generated such that the *SVD* components $u_i'^T b^{\rm exact}$ of the corresponding exact right-hand side $b^{\rm exact} = Ax^{\rm exact}$ appear as shown in Fig. 3. The scaling and the perturbations are similar to those for the *R-TLS* example above, except that the noise levels are smaller:

$$\sigma = 0.001,\ 0.002,\ 0.004,\ 0.008.$$

Plots of the relative errors $\|x^{\rm exact} - x_k\|_2/\|x^{\rm exact}\|_2$ and $\|x^{\rm exact} - \bar{x}_k\|_2/\|x^{\rm exact}\|_2$ for the *TSVD* and *T-TLS* methods are shown in Fig. 4. For small noise levels, the minimum

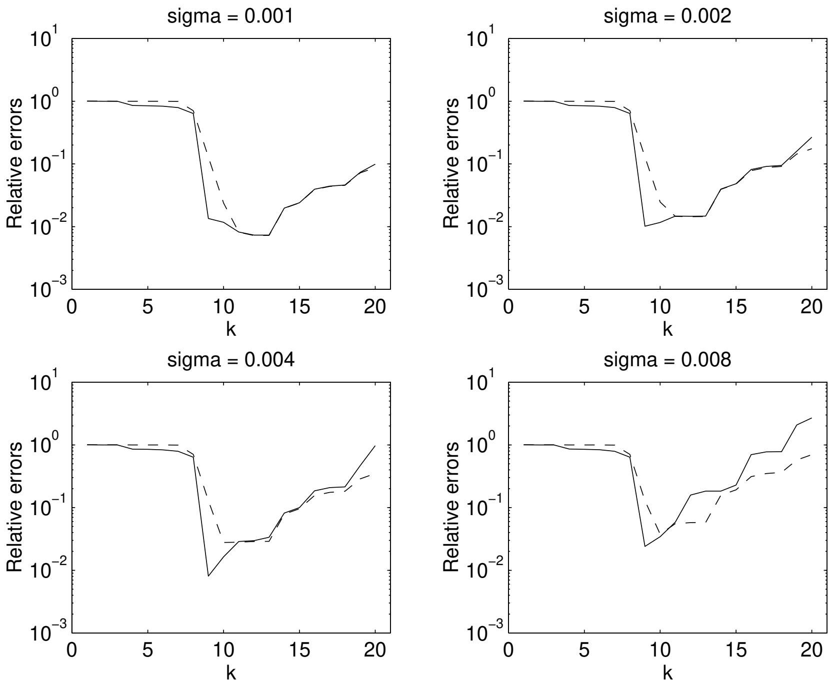

Fig. 4. *Plots of the relative errors in the TSVD solutions (dashed lines) and T-TLS solutions (solid lines) versus λ_L for four values of the noise level.*

relative errors are identical. As the noise level increases, the minimum *T-TLS* errors become smaller than those of the *TSVD* method. Histograms showing the results of many similar experiments can be found in [4], supporting the conclusion that *T-TLS* is superior to *TSVD*, as well as Tikhonov's method, whenever the noise level is large and the solution has large *SVD* components corresponding to the smallest retained singular values.

5 Conclusion

We have presented two different approaches to incorporation of regularization, or stabilization, into the *TLS* setting. The two methods are similar in spirit to Tikhonov regularization and truncated *SVD*, respectively. We have described the regularizing properties of the two methods and demonstrated by numerical examples that in certain cases with large perturbations, these new methods are able to yield more accurate regularized solutions than those produced by the standard methods.

References

[1] Å. Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[2] H. W. Engl, *Regularization methods for the stable solution of inverse problems*, Surv. Math. Ind. **3** (1993), pp. 71–143.

[3] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Vieweg, Wiesbaden, 1993.

[4] R. D. Fierro, G. H. Golub, P. C. Hansen and D. P. O'Leary, *Regularization by truncated total least squares*, Report UNIC-93-14 (20 pages), December 1993; revised July 1995; SIAM J. Sci. Comput., to appear.

[5] G. H. Golub, P. C. Hansen and D. P. O'Leary, *Tikhonov regularization and total least squares*, manuscript in preparation for SIAM J. Sci. Comput.

[6] M. Hanke and P. C. Hansen, *Regularization methods for large-scale problems*, Surv. Math. Ind. **3** (1993), pp. 253–315.

[7] P. C. Hansen, *Regularization Tools: A Matlab Package for Analysis and Solution of Discrete Ill-Posed Problems*; Numer. Algo. **6** (1994), pp. 1–35.

[8] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, Doctoral Dissertation, Polyteknisk Forlag, Denmark, 1996.

[9] T. A. Hua and R. F. Gunst, *Generalized ridge regression: a note on negative ridge parameters*, Commun. Statist. – Theor. Meth. **12** (1983), pp. 37–45.

[10] C. C. Paige and M. A. Saunders, *LSQR: an algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software **8** (1982), pp. 43–71.

[11] G. W. Stewart, *On the invariance of perturbed null vectors under column scaling*, Numer. Math. **44** (1984), pp. 61–65.

[12] S. Van Huffel & J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, 1991.

[13] S. Van Huffel, J. Vandewalle and A. Haegemans, *An efficient and reliable algorithm for computing the singular subspace of a matrix, associated with its smallest singular values*, J. Comp. Appl. Math. **19** (1987), pp 313–330.