TIKHONOV REGULARIZATION AND TOTAL LEAST SQUARES[*]

GENE H. GOLUB[†], PER CHRISTIAN HANSEN[‡], AND 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. We show how Tikhonov's regularization method, which in its original formulation involves a least squares problem, can be recast in a total least squares formulation suited for problems in which both the coefficient matrix and the right-hand side are known only approximately. We analyze the regularizing properties of this method and demonstrate by a numerical example that, in certain cases with large perturbations, the new method is superior to standard regularization methods.

Key words. total least squares, discrete ill-posed problems, regularization, bidiagonalization

AMS subject classifications. 65F20, 65F30

PII. S0895479897326432

1. Introduction. In this paper we study a class of methods for producing an approximate solution to a linear system of equations $A\,x \approx b$, where A is $m \times n$ with $m \geq n$. We assume that the elements of A and those of b are contaminated by some noise. An appropriate statement of the problem in the case of noisy data is the *total least squares* (TLS) formulation [5], [6, section 12.3], [17]:

$$(1) \qquad \min_{\tilde{A},\tilde{b},x} \|(A\,,\,b) - (\tilde{A}\,,\,\tilde{b})\|_F \qquad \text{subject to} \qquad \tilde{b} = \tilde{A}\,x.$$

In contrast, if the elements of A are exact and only b contains noise, then the corresponding formulation is the *least squares* (LS) problem

$$(2) \qquad \min_{x} \|Ax - b\|_2.$$

Our focus in this work is on very ill-conditioned problems where the singular values of A decay gradually to zero. Such problems arise, for example, from the discretization of ill-posed problems such as integral equations of the first kind. See, e.g., [2], [3], [7], [8], and [10] for examples and details.

In these problems, the solutions to the two formulations in (1) and (2) can be hopelessly contaminated by the noise in directions corresponding to small singular values of A or $(A\,,\,b)$. Because of this, it is necessary to compute a *regularized solution* in which the effect of such noise is filtered out. For LS problems, the filtering is often done by Tikhonov's regularization method, where a damping is added to each *singular value decomposition* (SVD) component of the solution, thus effectively filtering out

the components corresponding to the small singular values. Alternatively, one can truncate the SVD expansion of the solution, leaving out all the SVD components corresponding to the small singular values.

For the TLS problem, the truncation approach has already been studied by Fierro et al. [4]. In the present work we focus on the Tikhonov regularization approach for TLS. We thus arrive at a new regularization method in which stabilization enters the formulation in a natural way and which is able to produce regularized solutions with superior properties for certain problems in which the perturbations are large.

Our paper is organized as follows. In section 2 we introduce the regularized TLS method and we study its regularizing properties. Computational aspects are described in section 3, and we conclude the paper with a numerical example in section 4. A preliminary report on this work appeared as [12].

2. The regularized TLS method. Our regularization of the TLS problem is based on Tikhonov regularization. For the linear LS problem (2), a general version of Tikhonov's method [16] takes the form

$$(3) \qquad \min \left\{ \|A\,x - b\|_2^2 + \lambda\,\|L\,x\|_2^2 \right\},$$

where λ is a positive constant chosen to control the size of the solution vector and L is a matrix that defines a (semi)norm on the solution through which the "size" is measured [11]. Often, L represents the first or second derivative operator. If L is the identity matrix, then the Tikhonov problem is said to be in standard form. The solution x_λ to (3) solves the problem

$$(4) \qquad (A^T A + \lambda\,L^T L)x = A^T b.$$

As λ increases, the (semi)norm $\|L\,x\|_2$ of the solution vector decreases monotonically while the residual $\|A\,x - b\|_2$ increases monotonically.

Tikhonov regularization has an important equivalent formulation as

$$(5) \qquad \min \|A\,x - b\|_2 \qquad \text{subject to} \qquad \|L\,x\|_2 \leq \delta,$$

where δ is a positive constant. Problem (5) is an LS problem with a quadratic constraint and, using the Lagrange multiplier formulation,

$$(6) \qquad \mathcal{L}(x, \lambda) = \|A\,x - b\|_2^2 + \lambda\,\left(\|L\,x\|_2^2 - \delta^2\right),$$

it can be shown that if $\delta \leq \|x_{\mathrm{LS}}\|_2$, where x_{LS} is the LS solution to (2), then the solution x_δ to (5) is identical to the solution x_λ to (3) for an appropriately chosen λ, and there is a monotonic relation between the parameters δ and λ.

To carry this idea over to the TLS setting, we add an upper bound on the (semi)norm $\|L\,x\|_2$ of the solution vector x in the TLS problem (1). The formulation of the regularized TLS (R-TLS) problem thus becomes

$$(7) \qquad \min \|(A\,,\,b) - (\tilde{A}\,,\,\tilde{b})\|_F \qquad \text{subject to} \qquad \tilde{b} = \tilde{A}\,x, \quad \|L\,x\|_2 \leq \delta,$$

and the corresponding Lagrange multiplier formulation is

$$(8) \qquad \hat{\mathcal{L}}(\tilde{A}, x, \mu) = \|(A\,,\,b) - (\tilde{A}\,,\,\tilde{A}\,x)\|_F^2 + \mu\,(\|L\,x\|_2^2 - \delta^2),$$

where μ is the Lagrange multiplier, zero if the inequality constraint is inactive. The solution $\bar{x}_\delta$ to this problem is different from the solution x_{TLS} to (1) whenever δ is

less than $\|L\,x_{\mathrm{TLS}}\|_2$. The two solutions $\bar{x}_\delta$ and x_δ to the two regularized problems in (5) and (7) have a surprising relationship, explained by the following theorem.

THEOREM 2.1. *The R-TLS solution $\bar{x}_\delta$ to* (7), *with the inequality constraint replaced by equality, is a solution to the problem*

$$(9) \qquad (A^T A + \lambda_I I_n + \lambda_L L^T L)\,x = A^T b,$$

where the parameters λ_I and λ_L are given by

$$(10) \qquad \lambda_I = -\,\frac{\|b - A\,x\|_2^2}{1 + \|x\|_2^2},$$

$$(11) \qquad \lambda_L = \mu\,(1 + \|x\|_2^2)$$

and where μ is the Lagrange multiplier in (8). *The two parameters are related by*

$$(12) \qquad \lambda_L\,\delta^2 = b^T(b - A\,x) + \lambda_I.$$

Moreover, the TLS residual satisfies

$$(13) \qquad \|(A\,,\,b) - (\tilde{A}\,,\,\tilde{b})\|_F^2 = -\lambda_I.$$

Proof. Define $\tilde{r} \equiv b - \tilde{b} = b - \tilde{A}\,x$. We characterize the solution to (7) by setting the partial derivatives of the Lagrangian $\hat{\mathcal{L}}$ (8) to zero. Differentiation with respect to the entries in $\tilde{A}$ yields

$$\tilde{A} - A - \tilde{r}x^T = 0,$$

and differentiation with respect to the entries in x yields

$$(14) \qquad -\tilde{A}^T\tilde{r} + \mu\,L^T L\,x = 0$$

or

$$(\tilde{A}^T\tilde{A} + \mu\,L^T L)\,x = \tilde{A}^T b.$$

Since $A = \tilde{A} - \tilde{r}x^T$ and $\tilde{A}^T\tilde{r} = \mu\,L^T L\,x$, we see that

$$\begin{aligned} A^T A &= (\tilde{A}^T - x\tilde{r}^T)(\tilde{A} - \tilde{r}x^T) \\ &= \tilde{A}^T\tilde{A} - \mu\,xx^T L^T L + x\tilde{r}^T\tilde{r}x^T - \mu\,L^T L\,xx^T \end{aligned}$$

and

$$\tilde{A}^T b = A^T b + x\tilde{r}^T b.$$

Gathering terms, we arrive at (9), with

$$\begin{aligned} \lambda_I &= \mu\,\delta^2 - \|\tilde{r}\|_2^2\,\|x\|_2^2 - \tilde{r}^T b, \\ \lambda_L &= \mu\,(1 + \|x\|_2^2). \end{aligned}$$

To simplify the expression for λ_I, we first rewrite $\tilde{r}$ as

$$\tilde{r} = b - \tilde{A}\,x = b - (A + \tilde{r}\,x^T)\,x = b - A\,x - \tilde{r}\,x^T x,$$

from which we obtain the relation

$$\tilde{r}\,(1 + \|x\|_2^2) = b - A\,x \tag{15}$$

and, therefore, $(1 + \|x\|_2^2)\|\tilde{r}\|_2^2 = (1 + \|x\|_2^2)^{-1}\|b - A\,x\|_2^2$. In (14), multiplication from the left by x^T leads to

$$\mu = \frac{x^T \tilde{A}^T \tilde{r}}{x^T L^T L x} = \frac{b^T \tilde{r} - \|\tilde{r}\|_2^2}{\delta^2} . \tag{16}$$

Inserting (15) and (16) into the above expression for λ_I, we obtain (10). Equation (12) is proved by multiplying λ_L by δ^2 and inserting (15) and (16). Finally, to prove the expression for the TLS residual, we use the relation

$$(A\,,\,b) - (\tilde{A}\,,\,\tilde{b}) = (A - \tilde{A}\,,\,b - \tilde{b}) = (-\tilde{r}\,x^T\,,\,\tilde{r}) = -\tilde{r}\left(\begin{array}{c} x \\ -1 \end{array}\right)^T .$$

Taking the Frobenius norm and using (15), we obtain (13).   □

This theorem characterizes the solution to the problem in (7) whenever δ is small enough that the constraint on Lx is binding. When the solution $\bar{x}_\delta$ to (7) satisfies $\|L\bar{x}_\delta\|_2 < \delta$ (with $L = 0$ as a special case), then (7) is just the standard TLS problem.

Below, we discuss the implications of this theorem for both the standard-form case $L = I_n$ (the identity matrix) and the general-form case $L \neq I_n$, both of which are important in applications.

2.1. The standard-form case. In the standard-form case, the problems simplify considerably. The Tikhonov problem (5) becomes

$$\min \|A\,x - b\|_2 \qquad \text{subject to} \qquad \|x\|_2 \le \delta, \tag{17}$$

with solution x_δ satisfying

$$(A^T A + \lambda\,I_n)x_\delta = A^T b \tag{18}$$

for some value of $\lambda \ge 0$. Similarly, the R-TLS problem takes the form

$$\min \|(A\,,\,b) - (\tilde{A}\,,\,\tilde{b})\|_F \qquad \text{subject to} \qquad \tilde{b} = \tilde{A}\,x, \quad \|x\|_2 \le \delta, \tag{19}$$

and (9) says that the solution $\bar{x}_\delta$ satisfies

$$(A^T A + \lambda_{IL} I_n)\,x = A^T b \tag{20}$$

whenever $\|\bar{x}_\delta\|_2 > \delta$, with $\lambda_{IL} = \lambda_I + \lambda_L$. Clearly, these two solutions are closely related.

THEOREM 2.2. *Let $L = I_n$ and let $\bar{\sigma}_{n+1}$ denote the smallest singular value of $(A\,,\,b)$. Then, for each value of δ, the resulting solutions $\bar{x}_\delta$ and x_δ are related as in the following chart.*

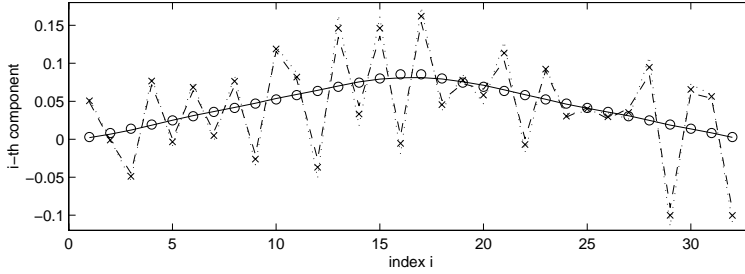| δ	Solutions	λ_{IL}
$\delta < \|x_{\mathrm{LS}}\|_2$	$\bar{x}_\delta = x_\delta$	$\lambda_{IL} > 0$
$\delta = \|x_{\mathrm{LS}}\|_2$	$\bar{x}_\delta = x_\delta = x_{\mathrm{LS}}$	$\lambda_{IL} = 0$
$\|x_{\mathrm{LS}}\|_2 < \delta < \|x_{\mathrm{TLS}}\|_2$	$\bar{x}_\delta \neq x_\delta = x_{\mathrm{LS}}$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

FIG. 1. *Regularized solutions to* (3) *for various values of the regularization parameter* λ; *see Table* 1 *for details.*

TABLE 1
The solutions shown in Figure 1.

Solution	λ	Solution norm	Symbol
x_λ	10^{-5}	0.289	solid line
x_{LS}	0	0.385	dashed line
x_λ	$-\frac{1}{2}\bar{\sigma}_{n+1}^2$	0.408	crosses
x_{TLS}	$-\bar{\sigma}_{n+1}^2$	0.443	dotted line
x^{exact}	$-$	0.288	circles

Proof. Consider the LS problem (17) first. If $\delta \geq \|x_{\text{LS}}\|$, then the constraint is not binding and the solution to the problem is x_{LS}, with $\lambda = 0$. As δ is reduced below $\|x_{\text{LS}}\|$, the value of λ increases monotonically, as can be seen by using the SVD $A = \Sigma_{i=1}^n u_i \sigma_i v_i^T$ in (18) to obtain

$$(21) \qquad x_\delta = \sum_{i=1}^n \frac{\sigma_i u_i^T b}{\sigma_i^2 + \lambda} v_i$$

and then taking norms.

Similarly, as long as $\delta \geq \|x_{\text{TLS}}\|$, the constraint in (19) is not binding, and in this case we have the standard TLS problem with $\lambda_{IL} = -\bar{\sigma}_{n+1}^2$ [17, Theorem 2.7]. As δ is reduced below $\|x_{\text{TLS}}\|$, (21) again shows that the λ_{IL} value must increase monotonically in order to reduce the solution norm.

Noting that $\|x_{\text{LS}}\| \leq \|x_{\text{TLS}}\|$ [17, Corollary 6.2], the theorem's conclusion follows.   □

In regularization problems, the norm $\|x_{\text{LS}}\|_2$ of the LS solution is normally very large due to the errors; thus, one will choose $\delta \leq \|x_{\text{LS}}\|_2$. When this is the case, we conclude from Theorem 2.2 that R-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 $\delta \leq \|x_{\text{LS}}\|_2$.

We remark that since $\|x_{\text{TLS}}\|_2 \geq \|x_{\text{LS}}\|_2$, there is usually a nontrivial set of "large" values of δ for which the multiplier λ_{IL} is negative. The corresponding R-TLS solutions $\bar{x}_\delta$ are distinctly different from the Tikhonov solutions and can be expected to be even more dominated by errors than the LS solution x_{LS}.

We illustrate Theorem 2.2 with an example: discretization of a Fredholm integral equation with the second derivative operator as kernel. The implementation is deriv2 from [9], the size of the matrix A is 64×32, and both A and b are perturbed by

Gaussian noise with zero mean and standard deviation 10^{-5}. We have $\bar{\sigma}_{n+1}^2 \approx 2.38 \cdot 10^{-9}$. Figure 1 shows the solutions listed in Table 1. We see that both x_{LS} and x_{TLS} have large oscillations (the oscillations of x_{TLS} being larger than those of x_{LS}), while the regularized LS solution x_λ corresponding to $\lambda = 10^{-5}$ is much smoother. The R-TLS solution, corresponding to $\lambda = -\frac{1}{2}\bar{\sigma}_{n+1}^2 = -1.19 \cdot 10^{-9}$, has large oscillations whose amplitude is between that of x_{TLS} and x_{LS}.

We conclude that if we use Tikhonov regularization with $L = I$ in order to reduce the norm of the solution below $\|x_{\text{LS}}\|_2$, then the R-TLS formulation produces the same solution as the regularized LS formulation. This is in contrast to the truncating approach, since we have shown in [4] that truncated TLS can be superior to truncated LS.

2.2. The general-form case. In many applications, it is necessary to choose a matrix L different from the identity matrix; these issues are discussed, e.g., in [10, section 4.3]. 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.

From Theorem 2.1 we notice that λ_L is always positive when $\delta < \|x_{\text{TLS}}\|_2$, because the Lagrange parameter μ is positive for these values of λ. On the other hand, λ_I is always negative and thus adds some *deregularization* to the solution. Statistical aspects of a negative regularization parameter in Tikhonov's method are discussed in [13].

For a given δ, there are usually several pairs of parameters λ_I and λ_L (and thus several solutions x) that satisfy relations (9)–(11), but only one of these solves the optimization problem (7). According to (13), this is the solution that corresponds to the smallest value of $|\lambda_I|$. The following relations hold.

THEOREM 2.3. *For a given value of δ, the solution $\bar{x}_\delta$ to the R-TLS problem is related to the solution to the TLS problem x_{TLS} as in the following chart.*

δ	Solution	λ_I		λ_L
$\delta < \|L\,x_{\text{TLS}}\|_2$	$\bar{x}_\delta \neq x_{\text{TLS}}$	$\lambda_I < 0$	$\partial\lambda_I/\partial\delta > 0$	$\lambda_L > 0$
$\delta \geq \|L\,x_{\text{TLS}}\|_2$	$\bar{x}_\delta = x_{\text{TLS}}$	$\lambda_I = -\bar{\sigma}_{n+1}^2$		$\lambda_L = 0$

Proof. For $\delta < \|L\,x_{\text{TLS}}\|_2$, the inequality constraint is binding and the Lagrange multiplier μ is positive, since this is a necessary condition for optimality; see, for example, [14]. Thus λ_L is positive. The residual (13) is monotonically decreasing as δ increases, since optimal solutions for smaller values of δ are candidate solutions for larger δs, so λ_I is monotonically increasing.

For $\delta = \|L\,x_{\text{TLS}}\|_2$, the Lagrange multiplier is zero, and the solution becomes the unconstrained minimizer x_{TLS}. The value $-\bar{\sigma}_{n+1}^2$ follows from [5, Theorem 4.1]. The constraint is never again binding for larger δ, so the solution to the problem remains unchanged. □

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 expression $\lambda \|L\,x\|_2$ in (3) is replaced with the norm $(\lambda_I\|x\|_2^2 + \lambda_L\|L\,x\|_2^2)^{1/2}$. If $\lambda_I I_n + \lambda_L L^T L$ is indefinite or negative definite then there is no equivalent interpretation.

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 (10) and is smallest in absolute value. The corresponding value of δ can then easily be computed from relation (12).

We now discuss how to solve (9) efficiently for many values of λ_I and λ_L. First, we note that the equation is equivalent to the augmented system

(22)
$$
\begin{pmatrix} I_m & 0 & A \\ 0 & I_p & \lambda_L^{1/2}L \\ A^T & \lambda_L^{1/2}L^T & -\lambda_I I_n \end{pmatrix} \begin{pmatrix} r \\ s \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}.
$$

Our algorithm is based on this formulation.

We assume that the matrix L is a banded matrix, often the case when L represents a derivative operator. The key to efficiency, then, is to reduce A to $n \times n$ bidiagonal form B by means of orthogonal transformations: $H^T A K = B$. The orthogonal right transformations also should 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 can recast the augmented system in (22) in the following form:

(23)
$$
\begin{pmatrix} I_n & 0 & B \\ 0 & I_p & \lambda_L^{1/2}C \\ B^T & \lambda_L^{1/2}C^T & -\lambda_I I_n \end{pmatrix} \begin{pmatrix} H^T r \\ J^T s \\ K^T x \end{pmatrix} = \begin{pmatrix} H^T b \\ 0 \\ 0 \end{pmatrix}.
$$

Since λ_I changes more frequently than λ_L in our approach, we will now use Givens rotations to annihilate $\lambda_L^{1/2}C$ using B by means of Elden's algorithm [1, section 5.3.4], which can be represented as

$$
\begin{pmatrix} B \\ \lambda_L^{1/2}C \end{pmatrix} = G \begin{pmatrix} \hat{B} \\ 0 \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} \hat{B} \\ 0 \end{pmatrix}.
$$

When we insert this G into the augmented system (23), it becomes

$$
\begin{pmatrix} I_n & 0 & \hat{B} \\ 0 & I_p & 0 \\ \hat{B}^T & 0 & -\lambda_I I_n \end{pmatrix} \begin{pmatrix} \hat{r} \\ \hat{s} \\ K^T x \end{pmatrix} = \begin{pmatrix} G_{11}^T H^T b \\ G_{12}^T H^T b \\ 0 \end{pmatrix}.
$$

The middle block row is now decoupled, and we obtain

$$
\begin{pmatrix} I_n & \hat{B} \\ \hat{B}^T & -\lambda_I I_n \end{pmatrix} \begin{pmatrix} \hat{r} \\ K^T x \end{pmatrix} = \begin{pmatrix} G_{11}^T H^T b \\ 0 \end{pmatrix}.
$$

Finally, we apply a symmetric perfect shuffle reordering

$$
n+1, 1, n+2, 2, n+3, 3, \ldots, n, 2n
$$

to the rows and columns of the above matrix, to obtain a symmetric, tridiagonal, indefinite matrix of size $2n \times 2n$:

$$
\begin{pmatrix} -\lambda_I & \hat{b}_{11} & & & \\ \hat{b}_{11} & 1 & \hat{b}_{12} & & \\ & \hat{b}_{12} & -\lambda_I & \hat{b}_{22} & \\ & & \hat{b}_{22} & 1 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}.
$$

Up until now, we have used only orthogonal transformations in order to preserve numerical stability. We can solve this permuted system in a stable way using a general tridiagonal solver (e.g., Gauss elimination with partial pivoting).

4. Numerical results. In this section we present a numerical example that demonstrates the usefulness of the R-TLS method. Our computations are carried out in MATLAB using the *Regularization Tools* package [9].

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 LS solution; see [15]. The same observation is made in [4] for the T-TLS solution, and the numerical results presented below also support this observation for the R-TLS method. We emphasize that the precise meaning of "small" depends on the particular problem.

In several test problems from [9], the R-TLS algorithm was able to compute better results than Tikhonov's method. The test problem we have chosen to illustrate the R-TLS algorithm is the one for which the improvement is most dramatic. This test problem is a discretization by means of Gauss–Laguerre quadrature of the inverse Laplace transform

$$\int_0^\infty \exp(-s\,t)\,f(t)\,dt = \frac{1}{2} - \frac{1}{s+4/25}, \quad 0 \le s,$$

$$f(t) = 1 - \exp(-t\,4/25),$$

originating from [18] and implemented in the function ilaplace(n,2) in [9], except that the constant $4/25$ was chosen instead of $1/2$. The matrix L approximates the first-derivative operator. The dimensions are $m = n = 16$, the matrix A and the exact solution x^* are scaled such that $\|A\|_F = \|A\,x^*\|_2 = 1$, and the perturbed right-hand side is generated as

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

where the elements of the perturbations E and e are from a normal distribution with zero mean and unit standard deviation.

Figure 2 shows the relative errors $\|x^* - x_\delta\|_2/\|x^*\|_2$ and $\|x^* - \bar{x}_\delta\|_2/\|x^*\|_2$ in the Tikhonov and R-TLS solutions, respectively, for four values of the noise level:

$$\sigma = 0.001,\ 0.01,\ 0.1,\ 1.$$

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 values 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 Figure 3 we have plotted the "optimal" Tikhonov and R-TLS solutions, defined as the solutions that correspond to the minima of the curves in Figure 2. In addition, we have plotted the exact solution x^*. Clearly, the addition of the term $\lambda_I I_n$ in (9) introduces a nonconstant component in the right part of the plot of the regularized solution; it is precisely this component that improves the R-TLS error compared to the Tikhonov error.

Conclusion. We have proved that the R-TLS solution to the regularized form (7) of the TLS problem is closely related to the Tikhonov solution to (3). In the standard-form case, $L = I_n$, the two solutions are identical, showing that Tikhonov
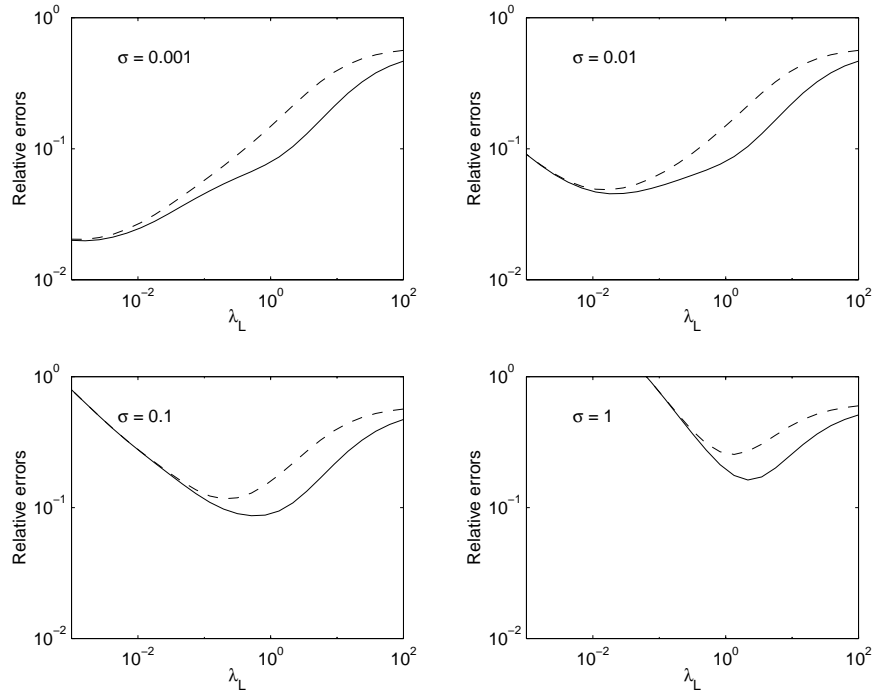
FIG. 2. *Plots of the relative errors in the Tikhonov solutions (dashed lines) and R-TLS solutions (solid lines) versus λ_L for four values of the noise level.*
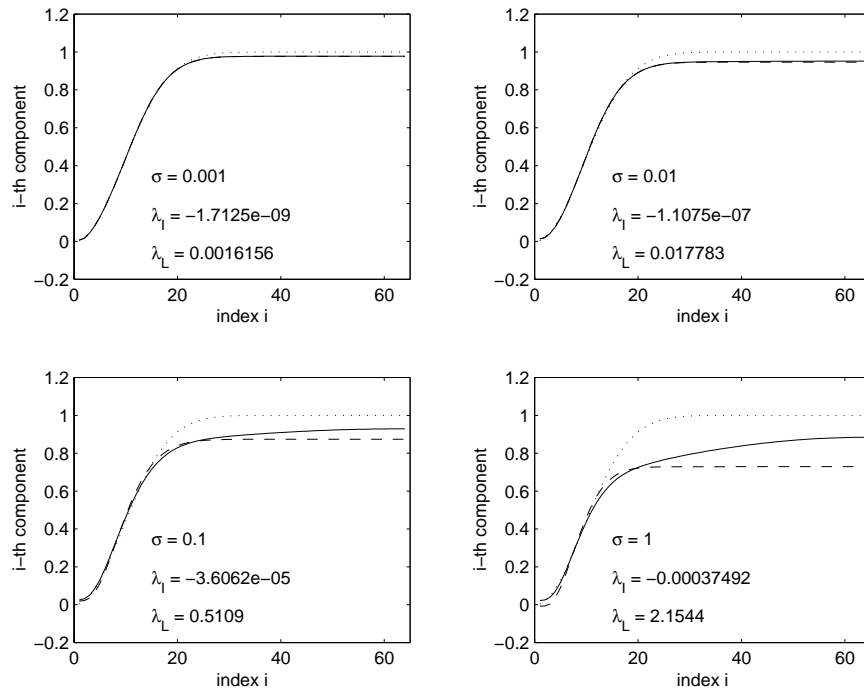

FIG. 3. *Plots of the "optimal" Tikhonov solutions (dashed lines) and R-TLS solutions (solid lines) for four values of the noise level. Also shown is the exact solution (dotted lines).*

regularization in this case is also suited for problems with a perturbed coefficient matrix. For general problems with $L \neq I_n$, the R-TLS solution is different from the corresponding Tikhonov solution in that a deregularizing term $\lambda_I I_n$ with a negative parameter λ_I is added to the formulation in (4).

REFERENCES

[1] A. BJÖRCK, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

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

[3] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, The Netherlands, 1996.

[4] R. D. FIERRO, G. H. GOLUB, P. C. HANSEN, AND D. P. O'LEARY, *Regularization by truncated total least squares*, SIAM J. Sci. Comput., 18 (1997), pp. 1223–1241.

[5] G. H. GOLUB AND C. F. VAN LOAN, *An analysis of the total least squares problem*, SIAM J. Numer. Anal., 17 (1980), pp. 883–893.

[6] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[7] C. W. GROETSCH, *Inverse Problems in the Mathematical Sciences*, Vieweg, Wiesbaden, Germany, 1993.

[8] M. HANKE AND P. C. HANSEN, *Regularization methods for large-scale problems*, Surveys Math. Indust., 3 (1993), pp. 253–315.

[9] P. C. HANSEN, *Regularization tools, a MATLAB package for analysis of discrete regularization problems*, Numer. Algorithms, 6 (1994), pp. 1–35.

[10] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.

[11] P. C. HANSEN AND D. P. O'LEARY, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.

[12] P. C. Hansen and D. P. O'Leary, *Regularization algorithms based on total least squares*, in Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling, S. Van Huffel, ed., SIAM, Philadelphia, 1997, pp. 127–137.

[13] T. A. HUA AND R. F. GUNST, *Generalized ridge regression: A note on negative ridge parameters*, Comm. Stat. Theory Methods, 12 (1983), pp. 37–45.

[14] S. G. NASH AND A. SOFER, *Linear and Nonlinear Programming*, McGraw–Hill, New York, 1996.

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

[16] A. N. TIKHONOV AND V. Y. ARSENIN, *Solutions of Ill-Posed Problems*, Wiley, New York, 1977.

[17] S. VAN HUFFEL AND J. VANDEWALLE, *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, 1991.

[18] J. M. VARAH, *On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems*, SIAM J. Numer. Anal., 10 (1973), pp. 257–267.