SIAM J. OPTIM.
Vol. 17, No. 1, pp. 98–118

ON THE SOLUTION OF THE TIKHONOV REGULARIZATION OF THE TOTAL LEAST SQUARES PROBLEM*

AMIR BECK[†] AND AHARON BEN-TAL[†]

Abstract. *Total least squares* (TLS) is a method for treating an overdetermined system of linear equations $\mathbf{Ax} \approx \mathbf{b}$, where both the matrix $\mathbf{A}$ and the vector $\mathbf{b}$ are contaminated by noise. *Tikhonov regularization* of the TLS (TRTLS) leads to an optimization problem of minimizing the sum of fractional quadratic and quadratic functions. As such, the problem is nonconvex. We show how to reduce the problem to a single variable minimization of a function $\mathcal{G}$ over a closed interval. Computing a value and a derivative of $\mathcal{G}$ consists of solving a single trust region subproblem. For the special case of regularization with a squared Euclidean norm we show that $\mathcal{G}$ is unimodal and provide an alternative algorithm, which requires only one spectral decomposition. A numerical example is given to illustrate the effectiveness of our method.

Key words. total least squares, Tikhonov regularization, fractional programming, nonconvex optimization, trust region subproblem

AMS subject classifications. 65F20, 90C20, 90C32

DOI. 10.1137/050624418

1. Introduction.

Many problems in data fitting and estimation give rise to an overdetermined system of linear equations $\mathbf{Ax} \approx \mathbf{b}$, where both the matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ and the vector $\mathbf{b} \in \mathbb{R}^m$ are contaminated by noise. The total least squares (TLS) approach to this problem [11, 12, 19] is to seek a perturbation matrix $\mathbf{E} \in \mathbb{R}^{m\times n}$ and a perturbation vector $\mathbf{r} \in \mathbb{R}^m$ that minimize $\|\mathbf{E}\|^2 + \|\mathbf{r}\|^2$ subject to the consistency equation $(\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b}+\mathbf{r}$ (here and elsewhere in this paper a matrix norm is always the Frobenius norm and a vector norm is the Euclidean one). The TLS approach was extensively used in a variety of scientific disciplines such as signal processing, automatic control, statistics, physics, economic, biology, and medicine (see, e.g., [19] and the references therein). The TLS problem has essentially an explicit solution, expressed by the singular value decomposition of the augmented matrix $(\mathbf{A}, \mathbf{b})$ (see, e.g., [11, 19]).

In practical situations, the original (noise-free) linear system is often ill-conditioned. For example, this happens when the system is obtained via discretization of ill-posed problems such as integral equations of the first kind (see, e.g., [10] and the references therein). In these cases the least squares (LS) solution as well as the TLS solution can be physically meaningless, and thus regularization is essential for stabilizing the solution.

There are two well-established approaches (among many others) to stabilize the LS solution: (i) *Tikhonov regularization*, where a quadratic penalty is appended to the LS objective function [4, 33], and (ii) *regularized least squares* (abbreviated RLS and LSQI), where a quadratic constraint bounding the size of the solution is added [4, 8].

*Received by the editors February 15, 2005; accepted for publication (in revised form) December 7, 2005; published electronically April 21, 2006.

http://www.siam.org/journals/siopt/17-1/62441.html

[†]MINERVA Optimization Center, Department of Industrial Engineering and Management Technion, Israel Institute of Technology, Haifa 3200, Israel (becka@ie.technion.ac.il, abental@ie.technion.ac.il). The research of the first author was partially supported by ISF grant 729/04. The research of the second author was partially supported by BSF grant 2002038.

For the TLS problem the situation is different. Stabilization by introducing a quadratic constraint was extensively studied [1, 10, 14, 28, 24]. On the other hand, Tikhonov regularization of the TLS (TRTLS) problem has not yet been considered.

In this paper we adopt the Tikhonov regularization concept to stabilize the TLS solution; i.e., we consider the problem

$$\text{(TRTLS)} \quad \min_{\mathbf{E},\mathbf{r},\mathbf{x}} \left\{ \|\mathbf{E}\|^2 + \|\mathbf{r}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 : (\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b}+\mathbf{r} \right\}, \tag{1}$$

where $\mathbf{L} \in \mathbb{R}^{k\times n}, k \le n$, is a full row rank matrix and $\rho > 0$ is a penalty parameter. $\mathbf{L}$ is a matrix that defines a (semi)norm on the solution through which its "size" is measured. A common example where $\mathbf{L}$ is not square is when $\mathbf{L}$ is an approximation matrix of the first or second order derivative [10, 16, 18].

The main difficulty associated with problem (TRTLS) is its nonconvexity. Nevertheless, we show in this paper that the problem can be solved efficiently to global optimality. First, in section 2 we reduce problem (TRTLS) to one involving only the $\mathbf{x}$ variables:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{\|\mathbf{x}\|^2+1} + \rho\|\mathbf{L}\mathbf{x}\|^2 \right\}. \tag{2}$$

In section 3 we derive an extremely mild condition for the attainability of an optimal solution to (2). An algorithm for solving problem (TRTLS) is then described in section 4. The algorithm consists of minimizing a single variable continuous (and differentiable under a mild condition) function $\mathcal{G}(\alpha)$ on a closed interval. Computing $\mathcal{G}(\alpha)$ and its derivative involves the solution of a single trust region subproblem. The interesting special case, where the matrix $\mathbf{L}$ in problem (TRTLS) is the identity matrix, is studied in section 5, where we prove that in this case $\mathcal{G}$ is unimodal and provide an alternative algorithm for solving the TRTLS problem requiring a single spectral decomposition. Finally, we provide in section 6 a detailed algorithm for the solution of the TRTLS problem (with a general regularization matrix) and demonstrate our method through an image deblurring example.

2. Simplified formulation of the TRTLS problem.

In order to simplify problem (1), we use a derivation similar to the one used in [1].[1] Problem (TRTLS) can be written as a double minimization problem:

$$\min_{\mathbf{x}} \min_{\mathbf{E},\mathbf{r}} \left\{ \|\mathbf{E}\|^2 + \|\mathbf{r}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 : (\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b}+\mathbf{r} \right\}. \tag{3}$$

Consider the inner minimization problem

$$\min_{\mathbf{E},\mathbf{r}} \left\{ \|\mathbf{E}\|^2 + \|\mathbf{r}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 : (\mathbf{A}+\mathbf{E})\mathbf{x} = \mathbf{b}+\mathbf{r} \right\}. \tag{4}$$

The Lagrangian of problem (4) is given by

$$\mathcal{L}(\mathbf{E},\mathbf{r},\boldsymbol{\lambda}) = \|\mathbf{E}\|^2 + \|\mathbf{r}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 + 2\boldsymbol{\lambda}^T((\mathbf{A}+\mathbf{E})\mathbf{x}-\mathbf{b}-\mathbf{r}).$$

Note that problem (4) is a linearly constrained convex problem with respect to the variables $\mathbf{E}$ and $\mathbf{r}$. Thus, the KKT conditions are necessary and sufficient [3, Proposition 3.4.1], and we conclude that $(\mathbf{E},\mathbf{r})$ is an optimal solution of (4) if and only if

[1] We thank Marc Teboulle for his contribution to this derivation.

there exists $\boldsymbol{\lambda} \in \mathbb{R}^m$ such that

$$2\mathbf{E} + 2\boldsymbol{\lambda}\mathbf{x}^T = \mathbf{0} \quad (\nabla_{\mathbf{E}}\mathcal{L} = 0), \tag{5}$$

$$2\mathbf{r} - 2\boldsymbol{\lambda} = \mathbf{0} \qquad (\nabla_{\mathbf{r}}\mathcal{L} = 0), \tag{6}$$

$$(\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{r} \quad \text{(feasibility)}. \tag{7}$$

From (6) we have $\boldsymbol{\lambda} = \mathbf{r}$. Substituting this into (5) we have

$$\mathbf{E} = -\mathbf{r}\mathbf{x}^T. \tag{8}$$

Combining (8) with (7) we obtain $(\mathbf{A} - \mathbf{r}\mathbf{x}^T)\mathbf{x} = \mathbf{b} + \mathbf{r}$, so

$$\mathbf{r} = \frac{\mathbf{A}\mathbf{x} - \mathbf{b}}{\|\mathbf{x}\|^2 + 1} \tag{9}$$

and consequently

$$\mathbf{E} = -\frac{(\mathbf{A}\mathbf{x} - \mathbf{b})\mathbf{x}^T}{\|\mathbf{x}\|^2 + 1}. \tag{10}$$

Finally, by substituting (9) and (10) into the objective function of problem (4) we obtain that the value of problem (4) is equal to $\frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{\|\mathbf{x}\|^2+1} + \rho\|\mathbf{L}\mathbf{x}\|^2$. Consequently, the TRTLS problem (1) reduces to

$$f^* = \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \mathcal{H}(\mathbf{x}) \equiv \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1} + \rho\|\mathbf{L}\mathbf{x}\|^2 \right\}. \tag{11}$$

For a given optimal solution $\mathbf{x}$ to the simplified TRTLS problem (11), the optimal pair $(\mathbf{E}, \mathbf{r})$ to the original TRTLS problem is given by (9) and (10).

3. Attainability of the minimum.

In this section, we find a sufficient condition for the attainability of the minimum in (11). First, notice that if $k = n$, then $\mathbf{L}$ has full rank and as a result the objective function is a coercive function[2] and the minimum is attained (see [3]). On the other hand, if $k < n$, then the minimum in (11) might not be attained. This is illustrated by the following example.

Example. Consider problem (11) with data

$$m = 3, \quad n = 2, \quad \mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{L} = \begin{pmatrix} 1 & 0 \end{pmatrix}, \ \rho = 1.$$

The TRTLS problem (11) in this case is

$$\min_{x_1, x_2} \left\{ \underbrace{\frac{(x_1 - 4)^2 + x_2^2}{1 + x_1^2 + x_2^2} + x_1^2}_{\mathcal{H}(x_1, x_2)} \right\}. \tag{12}$$

To show the nonattainment of the minimum, suppose on the contrary that the minimum is attained at a point (x_1^*, x_2^*). Notice that

$$(x_1^*)^2 \leq \mathcal{H}(x_1^*, x_2^*) \leq \mathcal{H}(0, x_2) \quad \forall x_2 \in \mathbb{R}.$$

[2] A real valued function $f : \mathbb{R}^n \to \mathbb{R}$ is coercive if $\lim_{\|\mathbf{x}\| \to \infty} f(\mathbf{x}) = \infty$.

Since $\mathcal{H}(0,x_2)=\frac{16+x_2^2}{1+x_2^2}\overset{x_2\to\infty}{\longrightarrow}1$ we conclude that $|x_1^*|\le 1$, which implies the inequality $(x_1^*-4)^2>1+(x_1^*)^2$. Therefore, the function $\varphi(y)=\mathcal{H}(x_1^*,y)=\frac{(x_1^*-4)^2+y^2}{1+(x_1^*)^2+y^2}+(x_1^*)^2$ is strictly decreasing and as a result we have, for example, $\mathcal{H}(x_1^*,x_2^*+1)<\mathcal{H}(x_1^*,x_2^*)$, which is a contradiction to the assumption that the minimum is attained at (x_1^*,x_2^*). We therefore conclude that the minimum (12) is not attained. □

Theorem 3.1 introduces a sufficient condition for the attainability of the minimum of the TRTLS problem (11).

THEOREM 3.1. *Consider problem* (11) *with* $\mathbf{A}\in\mathbb{R}^{m\times n},\mathbf{b}\in\mathbb{R}^m$, *and* $\mathbf{L}\in\mathbb{R}^{k\times n},n>k$. *Let* $\mathbf{F}\in\mathbb{R}^{n\times k}$ *be a matrix whose columns form an orthogonal basis for the null space of* $\mathbf{L}$. *If the following condition is satisfied,*

$$\lambda_{min}\begin{pmatrix}\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F} & \mathbf{F}^T\mathbf{A}^T\mathbf{b}\\ \mathbf{b}^T\mathbf{A}\mathbf{F} & \|\mathbf{b}\|^2\end{pmatrix}<\lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}),\tag{13}$$

then

(i)

$$f^*\le\lambda_{min}\begin{pmatrix}\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F} & \mathbf{F}^T\mathbf{A}^T\mathbf{b}\\ \mathbf{b}^T\mathbf{A}\mathbf{F} & \|\mathbf{b}\|^2\end{pmatrix};\tag{14}$$

(ii) *the minimum of* (11) *is attained.*

Proof. (i) Let $\mathbf{d}\in\mathbb{R}^{p+1}$ be an eigenvector corresponding to the minimum eigenvalue of the matrix

$$\mathbf{H}=\begin{pmatrix}\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F} & \mathbf{F}^T\mathbf{A}^T\mathbf{b}\\ \mathbf{b}^T\mathbf{A}\mathbf{F} & \|\mathbf{b}\|^2\end{pmatrix}.$$

Then

$$\frac{\mathbf{d}^T\mathbf{H}\mathbf{d}}{\|\mathbf{d}\|^2}=\lambda_{min}(\mathbf{H}).\tag{15}$$

d_{p+1} must be different from zero since otherwise we would have

$$\lambda_{min}(\mathbf{H})\overset{\mathbf{d}=(\tilde{\mathbf{d}}^T,0)^T}{=}\frac{\mathbf{d}^T\mathbf{H}\mathbf{d}}{\|\mathbf{d}\|^2}=\frac{\tilde{\mathbf{d}}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\tilde{\mathbf{d}}}{\|\tilde{\mathbf{d}}\|^2}\ge\lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}),$$

which is in contradiction to (13). Therefore, $d_{p+1}\neq 0$. Let $\mathbf{y}\in\mathbb{R}^p$ be such that $\frac{\mathbf{d}}{-d_{p+1}}=(\mathbf{y}^T,-1)^T$. Then

$$\begin{aligned}\lambda_{min}(\mathbf{H}) &\overset{(15)}{=} \frac{\mathbf{d}^T\mathbf{H}\mathbf{d}}{\|\mathbf{d}\|^2}=\frac{\left(\frac{\mathbf{d}}{d_{n+1}}\right)^T\mathbf{H}\left(\frac{\mathbf{d}}{d_{n+1}}\right)}{\left\|\left(\frac{\mathbf{d}}{d_{n+1}}\right)\right\|^2}\\ &= \frac{\begin{pmatrix}\mathbf{y}^T & -1\end{pmatrix}\mathbf{H}\begin{pmatrix}\mathbf{y}\\ -1\end{pmatrix}}{\|\begin{pmatrix}\mathbf{y}^T & -1\end{pmatrix}\|^2}=\frac{\mathbf{y}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{y}-2\mathbf{y}^T\mathbf{F}^T\mathbf{A}^T\mathbf{b}+\|\mathbf{b}\|^2}{\|\mathbf{y}\|^2+1}\\ &\overset{\mathbf{F}^T\mathbf{F}=\mathbf{I},\mathbf{L}\mathbf{F}=0}{=} \frac{\mathbf{y}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{y}-2\mathbf{y}^T\mathbf{F}^T\mathbf{A}^T\mathbf{b}+\|\mathbf{b}\|^2}{\mathbf{y}^T\mathbf{F}^T\mathbf{F}\mathbf{y}+1}+\rho\|\mathbf{L}\mathbf{F}\mathbf{y}\|^2\\ &= \mathcal{H}(\mathbf{F}\mathbf{y})\ge f^*,\end{aligned}$$

thus proving (i). To prove (ii), suppose on the contrary that the minimum value of (11), f^*, is not attained, which implies that there exists a sequence $\mathbf{x}_k, k \geq 1$, such that

$$\|\mathbf{x}_k\| \to \infty, \quad \underbrace{q(\mathbf{x}_k) + h(\mathbf{x}_k)}_{\mathcal{H}(\mathbf{x}_k)} \to f^*, \tag{16}$$

where $q(\mathbf{x}_k) \equiv \frac{\|\mathbf{A}\mathbf{x}_k - \mathbf{b}\|^2}{\|\mathbf{x}_k\|^2+1}$ and $h(\mathbf{x}_k) \equiv \rho\|\mathbf{L}\mathbf{x}_k\|^2$. Since both the sequences $q(\mathbf{x}_k)$ and $\frac{\mathbf{x}_k}{\|\mathbf{x}_k\|}$ are bounded, there exists a subsequence $\mathbf{x}_{n_k}$ for which the subsequences $q(\mathbf{x}_{n_k})$ and $\frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|}$ converge to a finite value. That is, there exist η and $\mathbf{d}$ such that

$$q(\mathbf{x}_{n_k}) \to \eta, \quad \frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|} \to \mathbf{d}.$$

Now, from (16) it follows that

$$\frac{q(\mathbf{x}_{n_k}) + h(\mathbf{x}_{n_k})}{\|\mathbf{x}_{n_k}\|^2} \to 0$$

and since $q(\mathbf{x}_{n_k})$ is bounded we have that $\frac{h(\mathbf{x}_{n_k})}{\|\mathbf{x}_{n_k}\|^2} \to 0$. But, on the other hand, $\frac{h(\mathbf{x}_{n_k})}{\|\mathbf{x}_{n_k}\|^2} \to \rho\|\mathbf{L}\mathbf{d}\|^2$ and as a result we have that $\|\mathbf{L}\mathbf{d}\|^2 = 0$, which is equivalent to $\mathbf{d} \in \text{Null}(\mathbf{L})$. To summarize, we have found a subsequence $\mathbf{x}_{n_k}$ for which $q(\mathbf{x}_{n_k})$ converges and $\frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|} \to \mathbf{d}$, where $\mathbf{d} \in \text{Null}(\mathbf{L})$ and $\|\mathbf{d}\| = 1$. Now,

$$\begin{aligned}
f^* &= \lim_{k\to\infty} \{q(\mathbf{x}_{n_k}) + h(\mathbf{x}_{n_k})\} \\
&\overset{h(\cdot)\geq 0}{\geq} \lim_{k\to\infty} q(\mathbf{x}_{n_k}) = \lim_{k\to\infty} \frac{\|\mathbf{A}\mathbf{x}_{n_k} - \mathbf{b}\|^2}{\|\mathbf{x}_{n_k}\|^2 + 1} \\
&= \lim_{k\to\infty} \frac{\mathbf{x}_{n_k}^T \mathbf{A}^T \mathbf{A}\mathbf{x}_{n_k} - 2\mathbf{b}^T\mathbf{A}\mathbf{x}_{n_k} + \|\mathbf{b}\|^2}{\|\mathbf{x}_{n_k}\|^2 + 1} \\
&= \lim_{k\to\infty} \frac{\left(\frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|}\right)^T \mathbf{A}^T\mathbf{A}\left(\frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|}\right) - 2\frac{1}{\|\mathbf{x}_{n_k}\|}\mathbf{b}^T\mathbf{A}\left(\frac{\mathbf{x}_{n_k}}{\|\mathbf{x}_{n_k}\|}\right) + \frac{\|\mathbf{b}\|^2}{\|\mathbf{x}_{n_k}\|^2}}{1 + \frac{1}{\|\mathbf{x}_{n_k}\|^2}} \\
&= \mathbf{d}^T\mathbf{A}^T\mathbf{A}\mathbf{d}.
\end{aligned}$$

Since $\mathbf{d} \in \text{Null}(\mathbf{L})$ we can write $\mathbf{d} = \mathbf{F}\mathbf{v}$, and therefore we obtain the following lower bound on f^*:

$$f^* \geq \min_{\mathbf{v}^T\mathbf{F}^T\mathbf{F}\mathbf{v}=1} \mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{v} \overset{\mathbf{F}^T\mathbf{F}=\mathbf{I}}{=} \min_{\|\mathbf{v}\|^2=1} \mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{v} = \lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}).$$

On the other hand, by condition (13), $\lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}^T) > \lambda_{min}(\mathbf{H})$, and therefore we have that

$$f^* > \lambda_{\min}(\mathbf{H}),$$

which is a contradiction to part (i). $\square$

Remarks.

1. Weak inequality is always satisfied in (13): the matrix in the right-hand side of (13) is a principal submatrix of the one in the left-hand side. Hence, by the interlacing theorem of eigenvalues [34, Theorem 7.8], weak inequality holds.
2. Condition (13) is invariant to the specific choice of the orthogonal basis of the null space of $\mathbf{L}$.
3. For $\mathbf{L} = \mathbf{0}$ problem (11) reduces to the classical TLS problem. In this case we can take $\mathbf{F} = \mathbf{I}$ in condition (13), which then reduces to the well-known condition [11, 19] for the attainability of the minimum in the TLS problem:

   $$\lambda_{min}\begin{pmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix} < \lambda_{min}\left(\mathbf{A}^T\mathbf{A}\right). \tag{17}$$

   Incidentally, for the nonregularized version of problem (12), i.e.,

   $$\min_{x_1,x_2}\left\{\frac{(x_1-4)^2+x_2^2}{1+x_1^2+x_2^2}\right\}, \tag{18}$$

   condition (17) does hold since

   $$\lambda_{min}\begin{pmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A} & \|\mathbf{b}\|^2 \end{pmatrix} = 0 < 1 = \lambda_{min}\left(\mathbf{A}^T\mathbf{A}\right)$$

   and indeed (18) attains an optimal solution $x_1^* = 4, x_2^* = 0$.
4. The TRTLS problem (12), for which nonattainability of the minimum was established, indeed does not satisfy condition (13). $\mathbf{F}$ can be chosen to be $\binom{0}{1}$, and we have

   $$\lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}) = 1$$

   and

   $$\lambda_{min}\begin{pmatrix} \mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F} & \mathbf{F}^T\mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A}\mathbf{F} & \|\mathbf{b}\|^2 \end{pmatrix} = \lambda_{min}\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = 1.$$

4. Solving the TRTLS problem with general L.

In this section we consider the TRTLS problem (11) with a full row rank $k \times n$ regularization matrix $\mathbf{L}$. We will assume that condition (13) is satisfied, and therefore the minimum is attained.

Problem (11) can be formulated as a double minimization problem in the following way:

$$\min_{\alpha\geq 1}\min_{\|\mathbf{x}\|^2=\alpha-1}\left\{\frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{\alpha}+\rho\|\mathbf{L}\mathbf{x}\|^2\right\},$$

which can be written as

$$\min_{\alpha\geq 1}\{\mathcal{G}(\alpha)\}, \tag{19}$$

where

$$\mathcal{G}(\alpha) \equiv \min_{\|\mathbf{x}\|^2=\alpha-1}\left\{\frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{\alpha}+\rho\|\mathbf{L}\mathbf{x}\|^2\right\}. \tag{20}$$

Calculating function values of $\mathcal{G}$ requires solving a minimization problem with a quadratic objective function and a norm equality constraint. In section 4.1 we briefly review known results on this problem including necessary and sufficient optimality conditions. In section 4.2 continuity and differentiability of $\mathcal{G}$ are established under standard second order sufficiency conditions. In section 4.3 an upper bound $\bar{\alpha}$ on the value of the optimal α is derived. Thus, the TRTLS problem (11) is reduced to a one dimensional minimization of $\mathcal{G}$ over a finite interval $[1, \bar{\alpha}]$.

4.1. Minimization of a quadratic function subject to a norm equality constraint. In this section we consider the minimization problem

$$\min_{\|\mathbf{x}\|^2=\beta} \left\{\mathbf{x}^T\mathbf{Q}\mathbf{x} - 2\mathbf{f}^T\mathbf{x} + c\right\}. \tag{21}$$

We do not assume that $\mathbf{Q}$ is positive semidefinite, and therefore the objective function need not be convex. Problem (21) is the well-known *trust region subproblem* (TRS); it has been extensively studied from both theoretical and algorithmic aspects [2, 5, 7, 23, 27, 31].[3] Necessary and sufficient conditions for a (global) solution of (21) are well established [5, 7, 32].

THEOREM 4.1 (see [5, 7, 32]). *Consider problem* (21) *with a symmetric matrix* $\mathbf{Q} \in \mathbb{R}^{n\times n}, \mathbf{f} \in \mathbb{R}^n, c \in \mathbb{R}, \beta \in \mathbb{R}^+$. *Then* $\mathbf{x}^*$ *is an optimal solution of* (21) *if and only if there exists* $\lambda^* \in \mathbb{R}$ *such that*

$$(\mathbf{Q} - \lambda^*\mathbf{I})\mathbf{x}^* = \mathbf{f}, \tag{22}$$

$$\|\mathbf{x}^*\|^2 = \beta, \tag{23}$$

$$\mathbf{Q} - \lambda^*\mathbf{I} \succeq \mathbf{0}. \tag{24}$$

Moreover, if $\mathbf{f} \notin$ *Null*$(\mathbf{Q} - \lambda_{\min}(\mathbf{Q})\mathbf{I})^{\perp}$, *then the solution of problem* (21) *is unique.*

Many algorithms have been suggested to solve the TRS. A solution based on the complete spectral decomposition can be found in [8]. For medium and large-scale problems the latter approach is not applicable. Thus, several methods have been devised for these scenarios [5, 7, 13, 23, 25, 30, 29].

4.2. Continuity and differentiability of $\mathcal{G}$.

4.2.1. Continuity. The continuity of $\mathcal{G}(\alpha)$ for $\alpha > 1$ follows from a theorem by Gauvin and Dubeau [9, Theorem 3.3] . The notation in [9] is quite different from the notation in this paper, and therefore we will present the three sufficient conditions for continuity of $\mathcal{G}$ at a point $\bar{\alpha}$ from [9] in our terminology (the quotation from [9] is in italic).

1. *The feasible set* $\{\mathbf{x} : \|\mathbf{x}\|^2 = \bar{\alpha} - 1\}$ *is nonempty.* This condition is naturally satisfied for $\bar{\alpha} > 1$.
2. *There exists* $\epsilon > 0$ *such that* $\bigcup_{|\alpha-\bar{\alpha}|<\epsilon}\{\mathbf{x} : \|\mathbf{x}\|^2 = \alpha - 1\}$ *is compact.* This is also true in our problem since the union is equal to $\{\mathbf{x} : \bar{\alpha} - 1 - \epsilon \leq \|\mathbf{x}\|^2 \leq \bar{\alpha} - 1 + \epsilon\}$, which is obviously compact.
3. *The Mangasarian–Fromovitz regularity conditions are satisfied (see* [22]*).* In our problem, this means that the gradient of the constraint is different from zero at the optimal solution, i.e., $\mathbf{x}^* \neq 0$. This is true for $\bar{\alpha} > 1$ since $\|\mathbf{x}^*\|^2 = \bar{\alpha} - 1$.

[3] The TRS is usually considered with an inequality constraint $\|\mathbf{x}\|^2 \leq \beta$ instead of an equality one; however, all known results can be trivially converted to the equality case.

What is left to prove is that $\mathcal{G}$ is continuous at $\alpha = 1$ (from the right). This is proved next.

LEMMA 4.1. *$\mathcal{G}$ is continuous at $\alpha = 1$ from the right.*

Proof. First, $\mathcal{G}(1) = \|\mathbf{b}\|^2$. Now, for every $\alpha > 1$ let $\mathbf{x}_\alpha$ be such that $\mathcal{H}(\mathbf{x}_\alpha) = \mathcal{G}(\alpha)$ and $\|\mathbf{x}_\alpha\|^2 = \alpha - 1$. Then

$$\begin{aligned}
|\mathcal{G}(\alpha) - \mathcal{G}(1)| &= |\mathcal{H}(\mathbf{x}_\alpha) - \|\mathbf{b}\|^2| \\
&= \left| \frac{\|\mathbf{A}\mathbf{x}_\alpha - \mathbf{b}\|^2}{\alpha} + \rho\|\mathbf{L}\mathbf{x}_\alpha\|^2 - \|\mathbf{b}\|^2 \right| \\
&= \left| \left(\frac{1}{\alpha} - 1\right)\|\mathbf{b}\|^2 + \frac{\mathbf{x}_\alpha^T\mathbf{A}^T\mathbf{A}\mathbf{x}_\alpha - 2\mathbf{b}^T\mathbf{A}\mathbf{x}_\alpha}{\alpha} + \rho\mathbf{x}_\alpha^T\mathbf{L}^T\mathbf{L}\mathbf{x}_\alpha \right| \\
&\le \left(1 - \frac{1}{\alpha}\right)\|\mathbf{b}\|^2 + \left(\frac{\lambda_{max}(\mathbf{A}^T\mathbf{A})}{\alpha} + \rho\lambda_{max}(\mathbf{L}^T\mathbf{L})\right)\|\mathbf{x}_\alpha\|^2 \\
&\quad + 2\frac{\|\mathbf{A}^T\mathbf{b}\|}{\alpha}\|\mathbf{x}_\alpha\| \\
&\overset{\|\mathbf{x}_\alpha\|^2=\alpha-1}{=} \left(1 - \frac{1}{\alpha}\right)\|\mathbf{b}\|^2 + \left(\frac{\lambda_{max}(\mathbf{A}^T\mathbf{A})}{\alpha} + \rho\lambda_{max}(\mathbf{L}^T\mathbf{L})\right)(\alpha - 1) \\
&\quad + 2\frac{\|\mathbf{A}^T\mathbf{b}\|}{\alpha}\sqrt{\alpha - 1} \\
&\xrightarrow{\alpha\to 1^+} 0.
\end{aligned}$$

Therefore, $\lim_{\alpha\to 1^+}\mathcal{G}(\alpha) = \mathcal{G}(1)$. □

COROLLARY 4.1. *$\mathcal{G}$ is continuous over* $[1, \infty)$.

4.2.2. Differentiability.

The function $\mathcal{G}$ is of the general form

$$\mathcal{G}(\alpha) = \min_{g(\mathbf{x})=\alpha-1} f(\mathbf{x}, \alpha), \tag{25}$$

where

$$f(\mathbf{x}, \alpha) \equiv \mathbf{x}^T\mathbf{Q}_\alpha\mathbf{x} - 2\mathbf{f}_\alpha^T\mathbf{x} + c_\alpha$$

and

$$g(\mathbf{x}) = \|\mathbf{x}\|^2,\ \mathbf{Q}_\alpha = \frac{1}{\alpha}\mathbf{A}^T\mathbf{A} + \rho\mathbf{L}^T\mathbf{L},\ \mathbf{f}_\alpha = \frac{1}{\alpha}\mathbf{A}^T\mathbf{b},\ c_\alpha = \frac{1}{\alpha}\|\mathbf{b}\|^2. \tag{26}$$

The single variable function $\mathcal{G}$ is not necessarily differentiable. In this subsection we show that under a suitable condition, $\mathcal{G}$ is differentiable of any order.

Our argument is the same as the one used in the sensitivity analysis of minimization problems (see, e.g., [3, 26] and the references therein). Theorem 4.2 establishes the differentiability of $\mathcal{G}$ under a suitable regularity condition.

THEOREM 4.2. *For every $\alpha > 1$ that satisfies the condition*

$$\mathbf{f}_\alpha \notin Null(\mathbf{Q}_\alpha - \lambda_{min}(\mathbf{Q}_\alpha)\mathbf{I})^\perp, \tag{27}$$

$\mathcal{G}(\alpha)$ is differentiable of any order and its first derivative is given by

$$\mathcal{G}'(\alpha) = \lambda(\alpha) + f'_\alpha(\mathbf{x}(\alpha), \alpha) = \lambda(\alpha) - \frac{\|\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b}\|^2}{\alpha^2}, \tag{28}$$

where $\mathbf{x}(\alpha)$ *and* $\lambda(\alpha)$ *are the unique solutions of the KKT conditions* (22) *and* (23).

Proof. Let $\alpha > 1$ be such that condition (27) is satisfied. By Theorem 4.1, condition (27) implies the uniqueness of the solution of the minimization problem (25). Consider the system of equations

$$(\mathbf{Q}_\alpha - \lambda\mathbf{I})\mathbf{x} = \mathbf{f}_\alpha, \tag{29}$$

$$\|\mathbf{x}\|^2 = \alpha - 1. \tag{30}$$

By Theorem 4.1, $x(\alpha)$ and $\lambda(\alpha)$ are the solutions of the system for the given α. The Jacobian matrix associated with the system of equations (29) and (30) with respect to $(\mathbf{x}, \lambda)$ at $(\mathbf{x}(\alpha), \lambda(\alpha))$ is given by

$$J = \begin{pmatrix} \mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I} & \mathbf{x}(\alpha) \\ \mathbf{x}(\alpha)^T & 0 \end{pmatrix}.$$

To show that J is nonsingular note first that condition (27) implies also that

$$\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I} \succ \mathbf{0}. \tag{31}$$

This is true since (29) implies that $\mathbf{f}_\alpha \in \text{Range}(\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I}) = \text{Null}(\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I})^\perp$. This condition combined with (27) and (24) implies that $\lambda(\alpha) < \lambda_{min}(\mathbf{Q}_\alpha)$. To show the nonsingularity of J, we will prove that the only solution of the system

$$J \begin{pmatrix} \mathbf{w} \\ t \end{pmatrix} = \mathbf{0}, \quad \mathbf{w} \in \mathbb{R}^n,\ t \in \mathbb{R},$$

is the trivial solution. Indeed, the system can be written explicitly as

$$(\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I})\mathbf{w} + 2t\mathbf{x}(\alpha) = \mathbf{0}, \tag{32}$$

$$2\mathbf{x}(\alpha)^T\mathbf{w} = \mathbf{0}. \tag{33}$$

Multiplying (32) by $\mathbf{w}^T$ from the left and using (33), we obtain $\mathbf{w}^T(\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I})\mathbf{w} = 0$. Since $\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I} \succ \mathbf{0}$ we conclude that $\mathbf{w} = 0$. Substituting this into (32) we have $t = 0$, proving the nonsingularity of J. Invoking the implicit function theorem, the differentiability of any order of $\mathbf{x}(\alpha)$ and $\lambda(\alpha)$ in a neighborhood of α follows. Now $\mathbf{x}(\alpha)$ and $\lambda(\alpha)$ satisfy the identities (in α)

$$f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) - \lambda(\alpha)g'_{\mathbf{x}}(\mathbf{x}(\alpha)) = 0, \tag{34}$$

$$g(\mathbf{x}(\alpha)) = \alpha - 1. \tag{35}$$

Differentiating both sides of (35) yields the equation

$$\dot{\mathbf{x}}(\alpha)^T g'_{\mathbf{x}}(\mathbf{x}(\alpha)) = 1. \tag{36}$$

Multiplying (34) from the left by $\dot{\mathbf{x}}(\alpha)^T$ we obtain

$$\dot{\mathbf{x}}(\alpha)^T f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) - \lambda(\alpha)\dot{\mathbf{x}}(\alpha)^T g'_{\mathbf{x}}(\mathbf{x}(\alpha)) = 0. \tag{37}$$

By substituting (36) into (37) we obtain

$$\dot{\mathbf{x}}(\alpha)^T f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) = \lambda(\alpha). \tag{38}$$

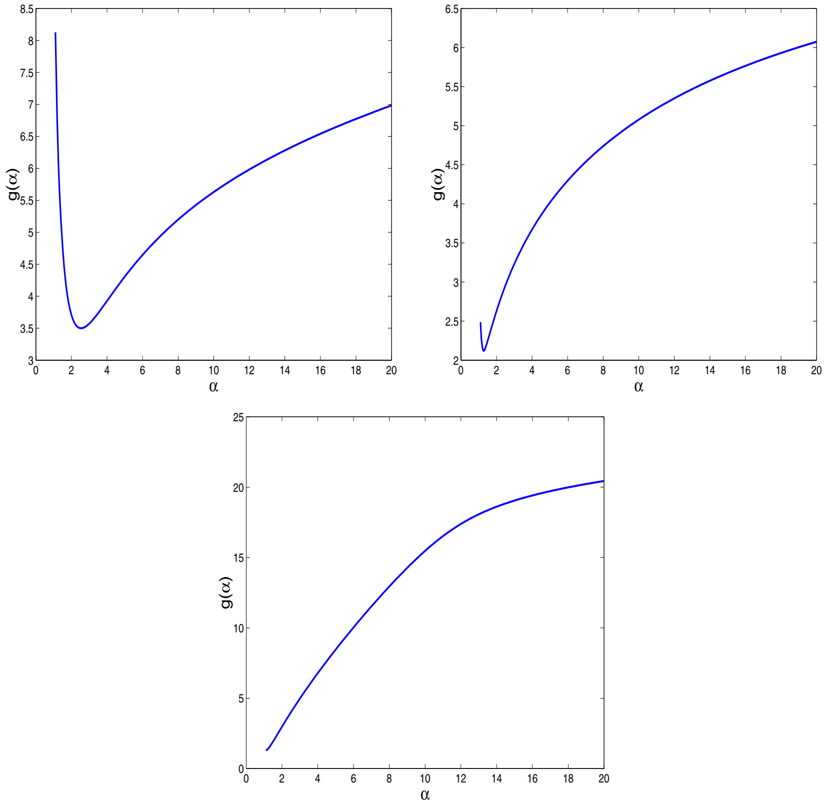


Fig. 1. *Examples of* $\mathcal{G}(\alpha)$.

$\mathcal{G}(\alpha)$ and its derivatives are given by

$$\begin{aligned} \mathcal{G}(\alpha) &= f(\mathbf{x}(\alpha), \alpha), \\ \mathcal{G}'(\alpha) &= \dot{\mathbf{x}}(\alpha)^T f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) + f'_{\alpha}(\mathbf{x}(\alpha), \alpha). \end{aligned} \tag{39}$$

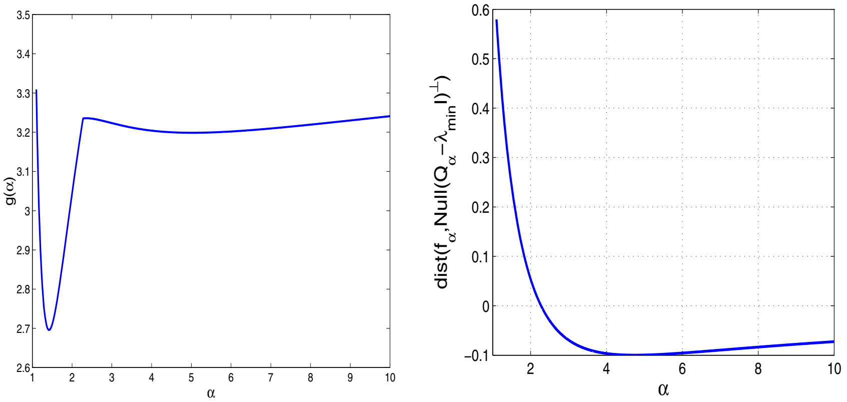
Substituting (38) into (39), the expression for the derivative (28) follows. □

Example. Some examples of $\mathcal{G}(\alpha)$ are given in Figure 1. These examples were randomly generated with dimensions $m = n = 4$ and $k = 3$.

In all of these examples $\mathcal{G}$ is continuous and differentiable. Note that in most examples the function $\mathcal{G}$ seems to be "well behaved" in the sense that it is strictly unimodal. A "bad" example is given in Figure 2(a), where we see an example of a nondifferentiable function. The point of nondifferentiability is $\bar{\alpha} = 2.275$. Figure 2(b) plots the quantity $dist(\mathbf{f}_\alpha, \mathrm{Null}(\mathbf{Q}_\alpha - \lambda_{min}(\mathbf{Q}_\alpha)\mathbf{I})^{\perp})$ versus α. It can be readily seen that the point in which the distance is zero is exactly the point $\bar{\alpha}$.

So far we have shown how to reduce the TRTLS problem (11) to a one dimensional problem $\min_{\alpha \geq 1} \mathcal{G}(\alpha)$. One of the problems frequently arising in one dimensional (line search) methods is determining an initial interval of search in which the optimum is known to reside. At this point, we have only shown that a lower bound on α is 1. Next we derive an upper bound.

4.3. Upper bound on the norm of optimal solutions. Let $\mathbf{x}^*$ be an optimal solution of problem (11). In this section we find an upper bound for $\|\mathbf{x}^*\|$. We recall



(a) (b)

Fig. 2. *An example of a nondifferentiable $\mathcal{G}(\alpha)$.*

the assumption that $\mathbf{L}$ is full row rank. In the case where $k = n$, it is very easy to bound the $\|\mathbf{x}^*\|$, as can be seen from the following lemma.

Lemma 4.2. *Suppose that $k = n$, and let $\mathbf{x}^*$ be an optimal solution of $\min_{\mathbf{x}\in\mathbb{R}^n} \mathcal{H}(\mathbf{x})$. Then $\|\mathbf{x}^*\|^2 \le \frac{\|\mathbf{b}\|^2}{\rho\lambda_{min}(\mathbf{L}\mathbf{L}^T)}$.*

Proof. First, notice that $\lambda_{min}(\mathbf{L}\mathbf{L}^T) > 0$ since $\mathbf{L}$ has full row rank. Now,

$$\rho\|\mathbf{L}\mathbf{x}^*\|^2 \le \mathcal{H}(\mathbf{x}^*) \le \mathcal{H}(0) = \|\mathbf{b}\|^2,$$

and the result follows from the simple observation that $\|\mathbf{L}\mathbf{x}^*\|^2 = (\mathbf{x}^*)^T\mathbf{L}^T\mathbf{L}\mathbf{x}^* \ge \lambda_{min}(\mathbf{L}\mathbf{L}^T)\|\mathbf{x}^*\|^2 > 0$. □

The case in which $k < n$ is much harder. In this case, we assume that condition (13) is satisfied.

Theorem 4.3. *Suppose that condition* (13) *is satisfied, and let $\mathbf{x}^*$ be an optimal solution of $\min_{\mathbf{x}\in\mathbb{R}^n} \mathcal{H}(\mathbf{x})$. Then*

$$\|\mathbf{x}^*\|^2 \le \max\left\{1, \frac{\|\mathbf{b}\|^2 + \lambda_{max}(\mathbf{A}^T\mathbf{A})(\delta+2\sqrt{\delta}) + \|\mathbf{A}^T\mathbf{b}\|(\delta+2\sqrt{\delta}) + l_1(1+\delta))}{l_1 - l_2}\right\}^2 + \delta, \tag{40}$$

where

$$l_2 = \lambda_{min}\begin{pmatrix} \mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F} & \mathbf{F}^T\mathbf{A}^T\mathbf{b} \\ \mathbf{b}^T\mathbf{A}\mathbf{F} & \|\mathbf{b}\|^2 \end{pmatrix},$$
$$l_1 = \lambda_{min}(\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}),$$

$\delta = \frac{l_2}{\lambda_{min}(\mathbf{L}\mathbf{L}^T)\rho}$*, and $\mathbf{F}$ is a matrix whose columns form an orthogonal base of Null$(\mathbf{L})$.*

Proof. Consider the decomposition

$$\mathbf{x}^* = \mathbf{F}\mathbf{v} + \mathbf{L}^T\mathbf{w}, \tag{41}$$

where $\mathbf{v} \in \mathbb{R}^{n-k}$ and $\mathbf{w} \in \mathbb{R}^n$ (such decomposition is possible since $\text{Null}(\mathbf{L}) = (\text{Range}(\mathbf{L}^T))^\perp$). Now,

$$\|\mathbf{x}^*\|^2 = \|\mathbf{v}\|^2 + \mathbf{w}^T\mathbf{L}\mathbf{L}^T\mathbf{w}. \tag{42}$$

By (14),

$$\mathcal{H}(\mathbf{x}^*) = f^* \leq l_2.$$

As a result,

$$\rho\|\mathbf{L}\mathbf{x}^*\|^2 \leq l_2. \tag{43}$$

Substituting (41) into (43) we obtain

$$\rho\mathbf{w}^T(\mathbf{L}\mathbf{L}^T)^2\mathbf{w} \leq l_2,$$

which implies the following inequality:

$$\mathbf{w}^T\mathbf{L}\mathbf{L}^T\mathbf{w} = \mathbf{w}^T(\mathbf{L}\mathbf{L}^T)^2\mathbf{w}\frac{\mathbf{w}^T\mathbf{L}\mathbf{L}^T\mathbf{w}}{\mathbf{w}^T(\mathbf{L}\mathbf{L}^T)^2\mathbf{w}} \leq \frac{l_2}{\rho}\lambda_{max}((\mathbf{L}\mathbf{L}^T)^{-1}(\mathbf{L}\mathbf{L}^T)(\mathbf{L}\mathbf{L}^T)^{-1}) = \delta. \tag{44}$$

We assume for now that $\|\mathbf{v}\| \geq 1$. Substituting the decomposition (41) into the objective function $\mathcal{H}$ we have

$$\begin{aligned}
&\mathcal{H}\left(\mathbf{x}^*\right)\\
&= \frac{\|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|^2}{\|\mathbf{x}^*\|^2+1} + \rho\|\mathbf{L}\mathbf{x}^*\|^2\\
&\geq \frac{\|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|^2}{\|\mathbf{x}^*\|^2+1} = \frac{\|\mathbf{A}(\mathbf{F}\mathbf{v}+\mathbf{L}^T\mathbf{w}) - \mathbf{b}\|^2}{\|\mathbf{F}\mathbf{v}+\mathbf{L}^T\mathbf{w}\|^2+1}\\
&= \frac{\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{v} + 2\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w} - 2\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{b} + \mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w} - 2\mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{b} + \|\mathbf{b}\|^2}{1+\|\mathbf{v}\|^2+\mathbf{w}^T\mathbf{L}\mathbf{L}^T\mathbf{w}}\\
&= \frac{\frac{\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{F}\mathbf{v}}{\|\mathbf{v}\|^2}+\beta}{1+\gamma} \geq \frac{l_1+\beta}{1+\gamma},
\end{aligned}$$

where

$$\begin{aligned}
\gamma &= \frac{1+\mathbf{w}^T\mathbf{L}\mathbf{L}^T\mathbf{w}}{\|\mathbf{v}\|^2},\\
\beta &= \frac{2\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w} - 2\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{b} + \mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w} - 2\mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{b} + \|\mathbf{b}\|^2}{\|\mathbf{v}\|^2}.
\end{aligned}$$

We have thus proven that $\mathcal{H}(\mathbf{x}^*) \geq \theta$, where $\theta = \frac{l_1+\beta}{1+\gamma}$. Combining this with Theorem 3.1 and condition (13) we have $\theta \leq l_2 < l_1$. Now,

$$l_1 - l_2 \leq l_1 - \theta = |\theta - l_1| = \left|\frac{l_1+\beta}{1+\gamma} - l_1\right| = \left|\frac{\beta - l_1\gamma}{1+\gamma}\right| \leq \beta + l_1\gamma. \tag{45}$$

Also,

$$\begin{aligned}
\gamma &\leq \frac{1+\delta}{\|\mathbf{v}\|^2} \overset{\|\mathbf{v}\|\geq 1}{\leq} \frac{1+\delta}{\|\mathbf{v}\|},\\
\beta &\leq \frac{2|\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w}| + 2|\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{b}| + |\mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w}| + 2|\mathbf{w}^T\mathbf{L}\mathbf{A}^T\mathbf{b}| + \|\mathbf{b}\|^2}{\|\mathbf{v}\|^2}\\
&\overset{(*)}{\leq} \frac{2}{\|\mathbf{v}\|}\left(\lambda_{max}(\mathbf{A}^T\mathbf{A})\sqrt{\delta} + \|\mathbf{A}^T\mathbf{b}\|\right) + \frac{1}{\|\mathbf{v}\|^2}\left(\|\mathbf{b}\|^2 + \lambda_{max}(\mathbf{A}^T\mathbf{A})\delta + 2\|\mathbf{A}^T\mathbf{b}\|\sqrt{\delta}\right)
\end{aligned} \tag{46}$$

$$
\begin{aligned}
&\overset{\|\mathbf{v}\|\geq 1}{\leq} \frac{2}{\|\mathbf{v}\|}\left(\lambda_{max}(\mathbf{A}^T\mathbf{A})\sqrt{\delta}+\|\mathbf{A}^T\mathbf{b}\|\right)+\frac{1}{\|\mathbf{v}\|}\left(\|\mathbf{b}\|^2+\lambda_{max}(\mathbf{A}^T\mathbf{A})\delta+2\|\mathbf{A}^T\mathbf{b}\|\sqrt{\delta}\right)\\
&= \frac{1}{\|\mathbf{v}\|}\left(\|\mathbf{b}\|^2+\lambda_{max}(\mathbf{A}^T\mathbf{A})(\delta+2\sqrt{\delta})+\|\mathbf{A}^T\mathbf{b}\|(2+2\sqrt{\delta})\right),
\end{aligned}
\tag{47}
$$

where inequality (*) is true due to the Cauchy–Schwarz inequality and trivial linear algebra inequalities. For example, $|\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w}|$ is bounded as follows:

$$
\begin{aligned}
|\mathbf{v}^T\mathbf{F}^T\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w}| &\overset{\text{C-S}}{\leq} \|\mathbf{F}\mathbf{v}\|\cdot\|\mathbf{A}^T\mathbf{A}\mathbf{L}^T\mathbf{w}\| \overset{\lambda_{max}(\mathbf{F})\leq 1}{\leq} \|\mathbf{v}\|\lambda_{max}(\mathbf{A}^T\mathbf{A})\|\mathbf{L}^T\mathbf{w}\|\\
&\overset{(44)}{\leq} \lambda_{max}(\mathbf{A}^T\mathbf{A})\|\mathbf{v}\|\sqrt{\delta}.
\end{aligned}
$$

Using the upper bound on β (47) and the upper bound on γ (46), we conclude that if $\|\mathbf{v}\|\geq 1$, then

$$
\begin{aligned}
l_1-l_2 &\overset{(45)}{\leq} \beta+l_1\gamma\\
&\leq \frac{1}{\|\mathbf{v}\|}\left(\|\mathbf{b}\|^2+\lambda_{max}(\mathbf{A}^T\mathbf{A})(\delta+2\sqrt{\delta})+\|\mathbf{A}^T\mathbf{b}\|(2+2\sqrt{\delta})+l_1(1+\delta)\right).
\end{aligned}
$$

Therefore,

$$
\|\mathbf{v}\|\leq\max\left\{1,\frac{\|\mathbf{b}\|^2+\lambda_{max}(\mathbf{A}^T\mathbf{A})(\delta+2\sqrt{\delta})+\|\mathbf{A}^T\mathbf{b}\|(\delta+2\sqrt{\delta})+l_1(1+\delta)}{l_1-l_2}\right\}.
\tag{48}
$$

Finally,

$$
\begin{aligned}
&\|\mathbf{x}^*\|^2\\
&= \|\mathbf{v}\|^2+\|\mathbf{L}^T\mathbf{w}\|^2\\
&\overset{(44),(48)}{\leq} \max\left\{1,\frac{\|\mathbf{b}\|^2+\lambda_{max}(\mathbf{A}^T\mathbf{A})(\delta+2\sqrt{\delta})+\|\mathbf{A}^T\mathbf{b}\|(\delta+2\sqrt{\delta})+l_1(1+\delta)}{l_1-l_2}\right\}^2+\delta. \qquad \square
\end{aligned}
$$

Remark. Recall that the sufficient condition for attainability is that $l_2 < l_1$. Note that if l_2 is very close to l_1, then the upper bound on $\|\mathbf{x}^*\|^2$ might be very large.

5. The case $\mathbf{L} = \mathbf{I}$.

5.1. Strict unimodality of $\mathcal{G}$.

In this section we show that in the case in which $\mathbf{L} = \mathbf{I}$, the function $\mathcal{G}$ defined in (20) has a very attractive property: *strictly unimodal*. A strictly unimodal function over an interval $[a, b]$ is a function that has a unique global minimum α^* and is strictly decreasing over $[a, \alpha^*]$ and strictly increasing over $[\alpha^*, b]$ (α^* can be equal to a or b and in that case the function is monotone). The fact that $\mathcal{G}$ is strictly unimodal implies that we can solve the one dimensional minimization problem efficiently (with, e.g., the golden section method; see [3]).

THEOREM 5.1. *Consider problem* (11) *with* $\mathbf{L} = \mathbf{I}$. *If* $\mathbf{A}^T\mathbf{b} \notin$ *Null*$(\mathbf{A}^T\mathbf{A} - \lambda_{\min}(\mathbf{A}^T\mathbf{A})\mathbf{I})^{\perp}$, *then* $\mathcal{G}$, *defined in* (20), *is differentiable for every* $\alpha > 1$ *and strictly unimodal.*

Proof. First, by substituting $\mathbf{Q}_\alpha = \frac{1}{\alpha}\mathbf{A}^T\mathbf{A} + \rho\mathbf{I}$ and $\mathbf{f}_\alpha = \frac{1}{\alpha}\mathbf{A}^T\mathbf{b}$ into (27) we obtain the following sufficient condition for differentiability of $\mathcal{G}$ at α:

$$
\mathbf{A}^T\mathbf{b} \notin \mathrm{Null}(\mathbf{A}^T\mathbf{A} - \lambda_{\min}(\mathbf{A}^T\mathbf{A})\mathbf{I})^{\perp}.
$$

Now, in order to prove the strict unimodality of $\mathcal{G}$, it is sufficient to prove the following property of $\mathcal{G}$: *if* $\mathcal{G}'(\alpha) = 0$, *then* $\mathcal{G}''(\alpha) > 0$. By differentiating both sides of (39), we obtain

$$\mathcal{G}''(\alpha) = \ddot{\mathbf{x}}(\alpha)^T f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) + \dot{\mathbf{x}}(\alpha)^T f''_{\mathbf{x}^2}(\mathbf{x}(\alpha), \alpha)\dot{\mathbf{x}}(\alpha) + 2\dot{\mathbf{x}}(\alpha)^T f''_{\mathbf{x}\alpha}(\mathbf{x}(\alpha), \alpha) + f''_{\alpha^2}(\mathbf{x}(\alpha), \alpha). \tag{49}$$

Differentiating (36), we have

$$\ddot{\mathbf{x}}(\alpha)^T g'_{\mathbf{x}}(\mathbf{x}(\alpha)) + \dot{\mathbf{x}}(\alpha)^T g''_{\mathbf{x}^2}(\mathbf{x}(\alpha))\dot{\mathbf{x}}(\alpha) = 0. \tag{50}$$

Therefore,

$$\begin{aligned}
\mathcal{G}''(\alpha) &= \mathcal{G}''(\alpha) - \lambda(\alpha) \cdot 0 \\
&\overset{(50)}{=} \mathcal{G}''(\alpha) - \lambda(\alpha)(\ddot{\mathbf{x}}(\alpha)^T g'_{\mathbf{x}}(\mathbf{x}(\alpha)) + \dot{\mathbf{x}}(\alpha)^T g''_{\mathbf{x}^2}(\mathbf{x}(\alpha))\dot{\mathbf{x}}(\alpha)) \\
&\overset{(49)}{=} \overbrace{\ddot{\mathbf{x}}(\alpha)^T \left(f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) - \lambda(\alpha) g'_{\mathbf{x}}(\mathbf{x}(\alpha))\right)}^{A} \\
&+ \overbrace{\dot{\mathbf{x}}(\alpha)^T \left(f''_{\mathbf{x}^2}(\mathbf{x}(\alpha), \alpha) - \lambda(\alpha) g''_{\mathbf{x}^2}(\mathbf{x}(\alpha))\right)\dot{\mathbf{x}}(\alpha)}^{B} \\
&+ \underbrace{2\dot{\mathbf{x}}(\alpha)^T f''_{\mathbf{x}\alpha}(\mathbf{x}(\alpha), \alpha) + f''_{\alpha^2}(\mathbf{x}(\alpha), \alpha)}_{C}.
\end{aligned}$$

By (34) we have $A = 0$ and

$$B = \dot{\mathbf{x}}(\alpha)^T \left(f''_{\mathbf{x}^2}(\mathbf{x}(\alpha), \alpha) - \lambda(\alpha) g''_{\mathbf{x}^2}(\mathbf{x}(\alpha))\right)\dot{\mathbf{x}}(\alpha) = \dot{\mathbf{x}}(\alpha)^T(\mathbf{Q}_\alpha - \lambda(\alpha)\mathbf{I})\dot{\mathbf{x}}(\alpha) \overset{(31)}{>} 0.$$

The latter inequality is true since by (36) $\dot{\mathbf{x}}(\alpha) \neq \mathbf{0}$. Suppose that $\mathcal{G}'(\alpha) = 0$; then

$$\dot{\mathbf{x}}(\alpha)^T f'_{\mathbf{x}}(\mathbf{x}(\alpha), \alpha) + f'_{\alpha}(\mathbf{x}(\alpha), \alpha) = 0,$$

which can also be written as

$$2\dot{\mathbf{x}}(\alpha)^T \left(\frac{\mathbf{A}^T(\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b})}{\alpha}\right) - \frac{\|\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b}\|^2}{\alpha^2} = -2\rho\dot{\mathbf{x}}(\alpha)^T \mathbf{L}^T \mathbf{L}\mathbf{x}(\alpha). \tag{51}$$

Now,

$$\begin{aligned}
C &= 2\dot{\mathbf{x}}(\alpha)^T f''_{\mathbf{x}\alpha}(\mathbf{x}(\alpha), \alpha) + f''_{\alpha^2}(\mathbf{x}(\alpha), \alpha) \\
&= -4\dot{\mathbf{x}}(\alpha)^T \frac{\mathbf{A}^T(\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b})}{\alpha^2} + 2\frac{\|\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b}\|^2}{\alpha^3} \\
&\overset{(51)}{=} 4\rho\frac{\dot{\mathbf{x}}(\alpha)^T \mathbf{L}^T \mathbf{L}\mathbf{x}(\alpha)}{\alpha}.
\end{aligned}$$

In our case $\mathbf{L} = \mathbf{I}$, and thus $C = 4\rho\frac{\dot{\mathbf{x}}(\alpha)^T \mathbf{x}(\alpha)}{\alpha} \overset{(36)}{=} \frac{2\rho}{\alpha} > 0$ and we conclude that, when $\mathcal{G}'(\alpha) = 0$, then $\mathcal{G}''(\alpha) = A + B + C > 0$, proving the unimodality property. ∎

5.2. Another approach to the case $\mathbf{L} = \mathbf{I}$.

5.2.1. The schematic algorithm. In the case $\mathbf{L} = \mathbf{I}$ the problem is given by

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \mathcal{H}(\mathbf{x}) \equiv \frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{\|\mathbf{x}\|^2+1} + \rho\|\mathbf{x}\|^2 \right\}. \tag{52}$$

We use the following simple observation, which goes back to Dinkelbach [6]: For every $t \in \mathbb{R}$, the following two statements are equivalent:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathcal{H}(\mathbf{x}) \le t,$$

$$\min_{\mathbf{x}\in\mathbb{R}^n} \{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \rho\|\mathbf{x}\|^4 + \rho\|\mathbf{x}\|^2 - t(\|\mathbf{x}\|^2+1)\} \le 0. \tag{53}$$

The minimization problem (53) also seems hard to solve; however, we will show in section 5.2.2 that it is in fact a very simple problem having essentially an explicit solution. Consider the function $\phi : \mathbb{R} \to \mathbb{R}$ defined by

$$\phi(t) = \min_{\mathbf{x}\in\mathbb{R}^n} \{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \rho\|\mathbf{x}\|^4 + \rho\|\mathbf{x}\|^2 - t(\|\mathbf{x}\|^2+1)\}.$$

We claim that ϕ is strictly decreasing. To prove this suppose that $t_1 < t_2$, and let $\mathbf{x}_{t_1} \equiv \operatorname{argmin}_{\mathbf{x}\in\mathbb{R}^n}\{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \rho\|\mathbf{x}\|^4 + \rho\|\mathbf{x}\|^2 - t_1(\|\mathbf{x}\|^2+1)\}$. Then

$$\begin{aligned} \phi(t_1) &= \|\mathbf{A}\mathbf{x}_{t_1}-\mathbf{b}\|^2 + \rho\|\mathbf{x}_{t_1}\|^4 + \rho\|\mathbf{x}_{t_1}\|^2 - t_1(\|\mathbf{x}_{t_1}\|^2+1) \\ &> \|\mathbf{A}\mathbf{x}_{t_1}-\mathbf{b}\|^2 + \rho\|\mathbf{x}_{t_1}\|^4 + \rho\|\mathbf{x}_{t_1}\|^2 - t_2(\|\mathbf{x}_{t_1}\|^2+1) \ge \phi(t_2). \end{aligned}$$

From the above observation we also have that $t^* \equiv \min_{\mathbf{x}\in\mathbb{R}^n} \mathcal{H}(\mathbf{x})$ is the unique root of $\phi(\cdot)$. Moreover, $t^* \in [0, \|\mathbf{b}\|^2]$ since

$$\phi(0) = \min_{\mathbf{x}\in\mathbb{R}^n} \{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \rho\|\mathbf{x}\|^4 + \rho\|\mathbf{x}\|^2\} \ge 0$$

and

$$\begin{aligned} \phi(\|\mathbf{b}\|^2) &= \min_{\mathbf{x}\in\mathbb{R}^n} \{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2 + \rho\|\mathbf{x}\|^4 + (\rho-\|\mathbf{b}\|^2)\|\mathbf{x}\|^2 - \|\mathbf{b}\|^2\} \\ &\le \min_{\mathbf{x}\in\mathbb{R}^n} \{\|\mathbf{A}\mathbf{0}-\mathbf{b}\|^2 + \rho\|\mathbf{0}\|^4 + (\rho-\|\mathbf{b}\|^2)\|\mathbf{0}\|^2 - \|\mathbf{b}\|^2\} = 0. \end{aligned}$$

As a result, the optimal t^* can be found by, e.g., a simple bisection algorithm with an initial interval $[0, \|\mathbf{b}\|^2]$.

5.2.2. Solving the subproblem. The subproblem can also be written as

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} + (\rho - t)\|\mathbf{x}\|^2 + \rho\|\mathbf{x}\|^4 - 2\mathbf{b}^T\mathbf{A}\mathbf{x} + \|\mathbf{b}\|^2 - t \right\}.$$

Making the change of variables $\mathbf{x} = \mathbf{U}\mathbf{z}$, where $\mathbf{U}$ is orthogonal matrix diagonalizing $\mathbf{A}^T\mathbf{A}$, i.e., $\mathbf{U}^T\mathbf{A}^T\mathbf{A}\mathbf{U} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, the problem then reduces to

$$\min_{\mathbf{z}\in\mathbb{R}^n} \sum_{j=1}^n \left\{ \lambda_j z_j^2 + (\rho - t)z_j^2 + \rho z_j^4 - 2f_j z_j \right\}, \tag{54}$$

where $\mathbf{f} = \mathbf{U}^T\mathbf{A}^T\mathbf{b}$. Note that since ρ canbe smaller than t, (54) might be a nonconvex problem. But, in fact, this does not really matter since this is a separable problem in its variables. Therefore, the solution of (54) requires solving n independent minimization problems:

$$\min_{\mathbf{z}_j\in\mathbb{R}} \left\{ (\lambda_j + \rho - t)z_j^2 + \rho z_j^4 - 2f_j z_j \right\}. \tag{55}$$

The scalar objective function is a coercive function (since the dominating factor is z_j^4). Therefore, the minimum of (55) is attained at a point satisfying $g_j'(z_j) = 0$, where $g_j(z_j) = (\lambda_j + \rho - t)z_j^2 + \rho z_j^4 - 2f_j z_j$. Therefore, the minimum is attained at one of the real roots of

$$4\rho z_j^3 + 2(\lambda_j + \rho - t)z_j - 2f_j = 0. \tag{56}$$

This is a cubic equation and therefore can be solved explicitly by Cardano's formula. More precisely, the roots of the cubic equation $x^3 + 3Qx - 2R = 0$ are given by

$$x_1 = (R + \sqrt{Q^3 + R^2})^{1/3} + (R - \sqrt{Q^3 + R^2})^{1/3}$$

and

$$\begin{aligned} x_{2,3} = &-\frac{1}{2}\left[(R + \sqrt{Q^3 + R^2})^{1/3} + (R - \sqrt{Q^3 + R^2})^{1/3}\right] \\ &\pm \frac{\sqrt{3}}{2} i \left[(R + \sqrt{Q^3 + R^2})^{1/3} - (R - \sqrt{Q^3 + R^2})^{1/3}\right]. \end{aligned}$$

In any case, it has three real roots if $Q^3 + R^2 \leq 0$ and only one real root (and two complex roots) otherwise. The minimum of (55) is attained at one of the roots of the cubic equation (56). Therefore, the initial step of the algorithm is to diagonalize the matrix $\mathbf{A}^T\mathbf{A}$, and then a bisection algorithm is invoked to find the unique root of the strictly decreasing function ϕ. The calculation of a function value of ϕ requires solving n cubic equations.

The algorithm described in this section is summarized below.

ALGORITHM TRTLSI.

```
Input: $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m, \rho > 0$, and $\epsilon$—a tolerance parameter.
Output: $\mathbf{x}^*$—a solution (up to some tolerance) of the TRTLS problem (11) with $\mathbf{L} = \mathbf{I}$.
    1. Set $t_{\min} \leftarrow 0$ and $t_{\max} \leftarrow \|\mathbf{b}\|^2$.
    2. Compute the spectral decomposition of $\mathbf{A}^T\mathbf{A}$: $\mathbf{U}^T\mathbf{A}^T\mathbf{A}\mathbf{U} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.
    3. Set $\mathbf{f} \leftarrow \mathbf{U}^T\mathbf{A}^T\mathbf{b}$.
    4. While $|t_{\max} - t_{\min}| > \epsilon$ repeat steps (a), (b), and (c):
        (a) For every $j = 1, 2, \ldots, n$ compute the solutions $z_1^j, \ldots, z_{p_j}^j$ of the one
            dimensional cubic equation (56). Here $p_j$ denotes the number of different
            real solutions of the $j$th cubic equation.
        (b) For every $j = 1, 2, \ldots, n$ set
                $\beta_j \leftarrow \min_{k=1,\ldots,p_j} \{(\lambda_j + \rho - t)(z_k^j)^2 + \rho(z_k^j)^4 - 2f_j z_k^j\}$.
        (c) If $\sum_{j=1}^n \beta_j - t < 0$, then $t_{\max} = t$; else $t_{\min} = t$.
    5. Set
                $m_j \leftarrow \operatorname{argmin}_{k=1,\ldots,p_j} \{(\lambda_j + \rho - t)(z_k^j)^2 + \rho(z_k^j)^4 - 2f_j z_k^j\}$.
    6. Let $\mathbf{w}$ be such that $w_j = z_{m_j}^j$ for every $j = 1, \ldots, n$.
    7. Set $\mathbf{x}^* = \mathbf{U}\mathbf{w}$.
```

The dominant computational effort when applying algorithm TRTLSI is the single calculation of the spectral decomposition of $\mathbf{A}^T\mathbf{A}$, which requires $O(n^3)$ operations. At each iteration the computational cost of solving n cubic equations is $O(n)$. For problems with up to several hundreds of variables, algorithm TRTLSI is therefore applicable. However, for problems with thousands or even tens of thousands of variables, algorithm TRTLSI cannot be implemented. Nevertheless, it is still possible to use the approach of solving the one dimensional minimization problem (19) since large-scale TRSs can be solved efficiently (see, e.g., [5, 7] and the references therein). A specific implementation of the algorithm for a general regularization matrix is given in the subsequent section.

6. Implementation and example.

We have shown that solving the TRTLS problem (11) (for a general regularization matrix $\mathbf{L}$) reduces to a problem of solving a one dimensional minimization problem over a closed interval. The specific details of the algorithm (for a general regularization matrix) depend on the choice of the one dimensional solver and the selection of a method for solving the TRS. In section 6.1 we describe a specific implementation—algorithm TRTLSG. We then apply the proposed algorithm in section 6.2 to an image deblurring example.

6.1. A detailed algorithm for the TRTLS problem.

We use the method of Moré and Sorensen for solving the TRS (21). The method is based on applying Newton's method to the problem

$$\frac{1}{\phi(\lambda)} - \frac{1}{\beta} = 0, \tag{57}$$

where $\phi(\lambda) \equiv \mathbf{f}^T(\mathbf{Q} - \lambda\mathbf{I})^{-1}\mathbf{f}$. The main computational effort at each iteration is the calculation of a Cholesky factorization of a matrix of the form $\mathbf{Q} - \lambda\mathbf{I}$. For large-scale problems the Cholesky factorization is not affordable, and other nondirect methods, such as Krylov subspace methods, can be employed (see, e.g., [29] and the references in [5, 7]). In our example $n = 1024$ so that Moré and Sorensen's method is appropriate.

```
ALGORITHM TRTLSG.
Input: $\mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m, \mathbf{L} \in \mathbb{R}^{k\times n}, \rho > 0$, and $\epsilon_1, \epsilon_2$—tolerance parameters.
Output: $\mathbf{x}^*$—a solution (up to some tolerance) of the TRTLS problem (11).
    1. Set $\alpha_{\min} \leftarrow 1 + \epsilon_1$.
    2. If $k = n$, set $\alpha_{\max}$ to be the upper bound given in Lemma 4.2; else $\alpha_{\max}$ is
       equal to the upper bound given in Theorem 4.3.
    3. While $|\alpha_{\max} - \alpha_{\min}| > \epsilon_2$ repeat steps (a), (b), and (c):
        (a) Set $\alpha \leftarrow \frac{\alpha_{min}+\alpha_{max}}{2}$.
        (b) Solve the following TRS:
                $\min_{\|\mathbf{x}\|^2=\alpha-1} \left\{\mathbf{x}^T\mathbf{Q}_\alpha\mathbf{x} - 2\mathbf{f}_\alpha^T\mathbf{x}\right\}$,
            where $\mathbf{Q}_\alpha$ and $\mathbf{f}_\alpha$ are given in (26), and obtain a solution $\mathbf{x}(\alpha)$ and a
            multiplier $\lambda(\alpha)$ that satisfy conditions (22), (23), and (24) (with $\mathbf{Q} =$
            $\mathbf{Q}_\alpha, \mathbf{f}_\alpha, \mathbf{x}^* = \mathbf{x}(\alpha)$, and $\lambda^* = \lambda(\alpha)$).
        (c) If $\underbrace{\lambda(\alpha) - \frac{\|\mathbf{A}\mathbf{x}(\alpha) - \mathbf{b}\|^2}{\alpha^2}}_{\mathcal{G}'(\alpha)} > 0$, then $\alpha_{\max} = \alpha$; else $\alpha_{\min} = \alpha$.
    4. Set $\mathbf{x}^* = \mathbf{x}(\alpha_{\max})$.
```

In our implementation the tolerance parameters take the values $\epsilon_1 = 10^{-1}$ and $\epsilon_2 = 10^{-6}$.

The one dimensional solver in algorithm TRTLSG is a simple bisection algorithm applied to the derivative of $\mathcal{G}(\alpha)$. To *guarantee* global convergence of the algorithm, the function $\mathcal{G}$ should be unimodal. For the case $\mathbf{L} = \mathbf{I}$ the unimodality property was proven in section 5.1. We observed through numerous random examples of the TRTLS problem of different dimensions ($4 \leq n, m, k \leq 1024$) that the unimodality property almost always holds even for $\mathbf{L} \neq \mathbf{I}$. The "bad" example in Figure 2 (with $m = n = 4, k = 3$) is an exceptional example. Moreover, for $n > 10$ we have not been able to find a single example which is not unimodal. Thus, for all practical purposes, algorithm TRTLSG finds the global optimum. If, for some reason, the function $\mathcal{G}$ is not unimodal, then algorithm TRTLSG does not necessarily converge to a global minimum and more sophisticated one dimensional global solvers should be employed.

6.2. Example. To illustrate the effectiveness of the TRTLS approach, we consider an image deblurring example. The TRTLS problems arising in this example were solved by algorithm TRTLSG implemented in MATLAB.

The choice of the regularization parameter ρ in our experiments was done by using the L-curve method [16, 21]. This method was originally devised as a method for choosing the regularization parameter for a regularized *least squares* problem. The L-curve is a plot of the L-norm $\|\mathbf{L}\mathbf{x}_\rho\|$ versus the residual $\|\mathbf{A}\mathbf{x}_\rho - \mathbf{b}\|$, where $\mathbf{x}_\rho$ is the solution of the regularization method with parameter ρ. The obtained plot usually has an L-shape appearance, and the chosen parameter is the one which is the closest to the left bottom corner. For the TLS problem, we follow the L-curve approach described in [24]: we plot the L-norm $\|\mathbf{L}\mathbf{x}_\rho\|^2$ versus the *fractional residual* $\|\mathbf{A}\mathbf{x}_\rho - \mathbf{b}\|^2/(1 + \|\mathbf{x}_\rho\|^2)$ for a various number of regularization parameters and pick the parameter closest to the L-shaped corner.

Let X be a 32×32 two dimensional image obtained from the sum of three harmonic oscillations:

$$X(z_1, z_2) = \sum_{l=1}^{3} a_i \cos(w_{l,1} z_1 + w_{l,2} z_2 + \phi_l), \quad \left(w_{l,i} = \frac{2\pi k_{l,i}}{n}\right), \; 1 \leq z_1, z_2 \leq 32,$$

where $k_{l,i} \in \mathbb{Z}^2$ (see Figure 3—true image). The specific values of the parameters are given in Table 1.

TABLE 1
Image parameters.

l	a_l	$w_{l,1}$	$w_{l,2}$	ϕ_l
1	1.3936	0.1473	0.0982	5.8777
2	0.5579	0.0982	0.0982	5.7611
3	0.8529	0.0491	0.0982	2.5778

We consider the square system

$$\mathbf{A}_{\text{true}} \mathbf{x}_{\text{true}} = \mathbf{b}_{\text{true}},$$

where $\mathbf{x}_{\text{true}} \in \mathbb{R}^{1024}$ is obtained by stacking the columns of the 32×32 image X. The vector $\mathbf{x}_{\text{true}}$ was normalized so that $\|\mathbf{x}_{\text{true}}\| = 1$. The 1024×1024 matrix $\mathbf{A}_{\text{true}}$ represents an atmospheric turbulence blur originating from [15] and implemented in the function blur(n,3) from the "Regularization Tools" [17]. The observed matrix

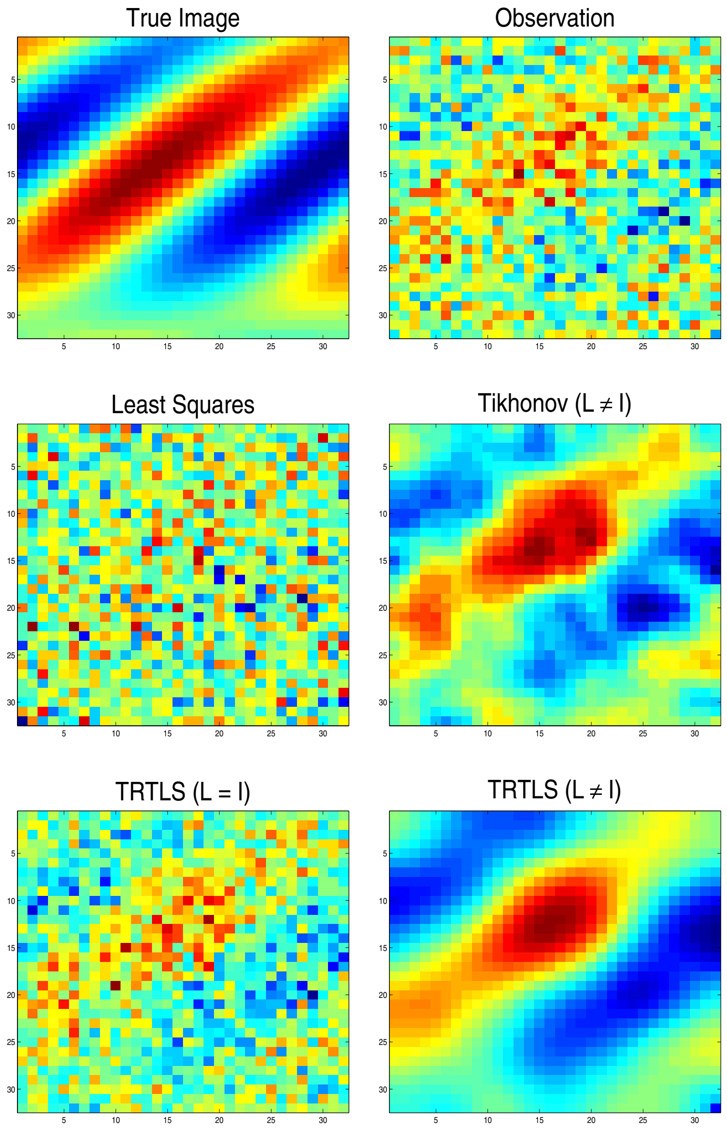


FIG. 3. *Results for different regularization solvers.*

and vector were generated by adding white noise to the data: $\mathbf{A} = \mathbf{A}_{true} + \sigma \mathbf{E}$ and $\mathbf{b} = \mathbf{b}_{\text{true}} + \sigma \mathbf{e}$, where each component of $\mathbf{E} \in \mathbb{R}^{1024\times 1024}$ and $\mathbf{e} \in \mathbb{R}^{1024}$ was generated from a standard normal distribution.

In our experiment the standard deviation σ was chosen to be 0.05, which results in a highly noisy image (see Figure 3—observation). The LS estimator was implemented in the function lsqr from [17]; it can be readily observed that it produces a poor image.

The choice of regularization matrix has a major influence on the quality of the obtained image. The solution of the TRTLS problem with *standard regularization* produces an unsatisfactory image (see Figure 3—TRTLS with $L = I$).

To produce a better result, we use a regularization matrix that accounts for the smoothness property of this image. In particular, the matrix $\mathbf{L}$ was chosen to satisfy the relation

$$\mathbf{L}^T\mathbf{L} = \mathbf{R}^T\mathbf{R} + \mathbf{I}, \tag{58}$$

where $\mathbf{R}$ is a discrete approximation of the Laplace operator, which is a two dimensional convolution with the following mask:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

This operator is standard in image processing [20]. With this choice of $\mathbf{L}$, the TRTLS algorithm gave the much better image (see Figure 3—TRTLS with $L \neq I$). We also compared our results to the one obtained by the classic Tikhonov regularization of the least squares, i.e., the solution of the minimization problem

$$\min_{\mathbf{x}}\{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2\}$$

with the same regularization matrix given in (58). Tikhonov regularization of the *least squares* (see Figure 3—Tikhonov $\mathbf{L} \neq \mathbf{I}$) provides a better image than the least squares, but its quality is inferior to the one obtained by the corresponding TRTLSG algorithm.

Acknowledgment. We give special thanks to the referees for their constructive comments and suggestions.

REFERENCES

[1] A. Beck, A. Ben-Tal, and M. Teboulle, *Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares*, SIAM J. Matrix Anal. Appl., to appear.

[2] A. Ben-Tal and M. Teboulle, *Hidden convexity in some nonconvex quadratically constrained quadratic programming*, Math. Programming, 72 (1996), pp. 51–63.

[3] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, Belmont, MA, 1999.

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

[5] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, MPS/SIAM Ser. Optim. 1, SIAM, Philadelphia, 2000.

[6] W. Dinkelbach, *On nonlinear fractional programming*, Management Sci., 13 (1967), pp. 492–498.

[7] C. Fortin and H. Wolkowicz, *The trust region subproblem and semidefinite programming*, Optim. Methods Softw., 19 (2004), pp. 41–67.

[8] W. Gander, G. H. Golub, and U. von Matt, *A constrained eigenvalue problem*, Linear Algebra Appl., 114/115 (1989), pp. 815–839.

[9] J. Gauvin and F. Dubeau, *Differential properties of the marginal function in mathematical programming*, Math. Programming Stud., 19 (1982), pp. 101–119.

[10] G. H. Golub, P. C. Hansen, and D. P. O'Leary, *Tikhonov regularization and total least squares*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 185–194.

[11] 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.

[12] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.

[13] N. I. M. Gould, S. Lucidi, M. Roma, and P. L. Toint, *Solving the trust-region subproblem using the Lanczos method*, SIAM J. Optim., 9 (1999), pp. 504–525.

[14] H. Guo and R. Renaut, *A regularized total least squares algorithm*, in Total Least Squares and Errors-in-Variables Modeling, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002, pp. 57–66.

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

[16] 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.

[17] P. C. Hansen, *Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems*, Numer. Algorithms, 6 (1994), pp. 1–35.

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

[19] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, Frontiers Appl. Math. 9, SIAM, Philadelphia, PA, 1991.

[20] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice–Hall, Englewood Cliffs, NJ, 1989.

[21] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, Prentice–Hall, Englewood Cliffs, NJ, 1974.

[22] O. L. Mangasarian, *Nonlinear programming*, McGraw–Hill, New York, 1969.

[23] J. J. Moré and D. C. Sorensen, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 553–572.

[24] R. A. Renaut and H. Guo, *Efficient algorithms for solution of regularized total least squares*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 457–476.

[25] F. Rendel and H. Wolkowicz, *A semidefinite framework for trust region subproblems with applications to large scale minimization*, Math. Programming, 77 (1997), pp. 273–299.

[26] A. Shapiro, *Second order sensitivity analysis and asymptotic theory of parameterized nonlinear programs*, Math. Programming, 33 (1985), pp. 280–299.

[27] B. W. Silverman, *On the estimation of a probability density function by the maximum penalized likelihood method*, Ann. Statist., 10 (1982), pp. 795–810.

[28] D. Sima, S. Van Huffel, and G. H. Golub, *Regularized total least squares based on quadratic eigenvalue problem solvers*, BIT, 44 (2004), pp. 793–812.

[29] D. C. Sorensen, *Minimization of a large-scale quadratic function subject to a spherical constraint*, SIAM J. Optim., 7 (1997), pp. 141–161.

[30] T. Steihaug, *The conjugate gradient method and trust regions in large scale optimization*, SIAM J. Numer. Anal., 20 (1983), pp. 626–637.

[31] R. J. Stern and H. Wolkowicz, *Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations*, SIAM J. Optim., 5 (1995), pp. 286–313.

[32] P. D. Tao and L. T. H. An, *A D.C. optimization algorithm for solving the trust-region subproblem*, SIAM J. Optim., 8 (1998), pp. 476–505.

[33] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, V.H. Winston, Washington, DC, 1977.

[34] F. Zhang, *Matrix Theory*, Springer, New York, 1999.