A CONTRIBUTION TO THE CONDITIONING OF THE TOTAL LEAST-SQUARES PROBLEM*

MARC BABOULIN[†] AND SERGE GRATTON[‡]

Abstract. We derive closed formulas for the condition number of a linear function of the total least-squares solution. Given an overdetermined linear system $Ax \approx b$, we show that this condition number can be computed using the singular values and the right singular vectors of $[A, b]$ and A. We also provide an upper bound that requires the computation of the largest and the smallest singular value of $[A, b]$ and the smallest singular value of A. In numerical examples, we compare these values and the resulting forward error bounds with the error estimates given by Van Huffel and Vandewalle [*The Total Least Squares Problem: Computational Aspects and Analysis*, Frontiers Appl. Math. 9, SIAM, Philadelphia, 1991], and we show the limitation of the first order approach.

1. Introduction. Given a matrix $A \in \mathbb{R}^{m \times n}$ $(m > n)$ and an observation vector $b \in \mathbb{R}^m$, the standard overdetermined linear least-squares (LS) problem consists in finding a vector $x \in \mathbb{R}^n$ such that Ax is the best approximation of b. Such a problem can be formulated using what is referred to as the linear statistical model

$$b = Ax + \epsilon, \quad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad \text{rank}(A) = n,$$

where ϵ is a vector of random errors having expected value $E(\epsilon) = 0$ and variance-covariance $V(\epsilon) = \sigma^2 I$.

In the linear statistical model, random errors affect exclusively the observation vector b, while A is considered as known exactly. However, it is often more realistic to consider that measurement errors might also affect A. This case is treated by the statistical model referred to as the errors-in-variables model (see, e.g., [17, p. 230] and [5, p. 176]), where we have the relation

$$(A + E)x = b + \epsilon.$$

In general it is assumed in this model that the rows of $[E, \epsilon]$ are independently and identically distributed with common zero mean vector and common covariance matrix. The corresponding linear algebra problem, discussed originally in [12], is called the total least-squares (TLS) problem and can be expressed as

$$(1.1) \qquad \min_{E, \epsilon} \|(E, \epsilon)\|_F, \qquad (A + E)x = b + \epsilon,$$

where $\| \cdot \|_F$ denotes the Frobenius matrix norm. As mentioned in [17, p. 238], the TLS method enables us to obtain a more accurate solution when entries of A are perturbed under certain conditions.

In error analysis, condition numbers are considered fundamental tools since they measure the effect on the solution of small changes in the data. In particular the conditioning of the LS problem was extensively studied in the numerical linear algebra literature (see, e.g., [5], [7], [8], [9], [10], [15], [16], [18], [21], [25]). The more general case of the conditioning of a linear function of an LS solution was studied in [2] and [4] when the perturbations of the data were measured, respectively, normwise and componentwise (note that the componentwise and normwise condition numbers for LS problems were also treated in [9] but without the generalization to a linear function of the solution). Moreover we can find in [3] algorithms using the software libraries LAPACK [1] and ScaLAPACK [6] as well as physical applications.

The notion of TLS was initially defined in the seminal paper [12] that was the first to propose a numerically stable algorithm. Then various aspects of the TLS problem were developed in the comprehensive book [17] including a large survey of theoretical bases and computational methods and applications, but also sensitivity analysis with, for instance, upper bounds for the TLS perturbation. The so-called scaled total least-squares (STLS) problem ($\min_{E,\epsilon}\|(E,\epsilon)\|_F$, $(A + E)x\gamma = \gamma b + \epsilon$ for a given scaling parameter γ) was formulated in [22] in which were addressed the difficulties coming from nonexistence of the TLS solution. In a recent paper [26], we can find sharp estimates of the normwise, mixed, and componentwise condition numbers of the STLS problem.

Here we are concerned with the TLS problem, which is a special case of the STLS problem, and we will consider perturbations of the data (A, b) that are measured normwise using a product norm. Contrary to [26], we will consider the general case of the conditioning of $L^T x$, a linear function of the TLS solution for which we will derive an exact formula. Considering the conditioning of $L^T x$ is relevant for many physical applications when the parameters to be estimated contain variables of physical significance as well as auxilliary variables (see, for example, the determination of GPS positions where the three-dimensional coordinates are the quantities of interest but the statistical model involves other parameters such as clock drift and GPS ambiguities [19] that are generally estimated during the solution process). Another example is related to the computation of the gravity field parameters, where regularization techniques are applied to some parameters which are more sensitive to perturbations [20].

The common situations correspond to the special cases where L is the identity matrix (condition number of the TLS solution) or a canonical vector (condition number of one solution component). The conditioning of a nonlinear function of a TLS solution can also be obtained by replacing, in the condition number expression, the quantity L^T by the Jacobian matrix at the solution.

We notice that the expression for the normwise condition number proposed in [26] is based on the evaluation of the norm of a matrix expressed as a Kronecker product, resulting in large matrices which may be, as pointed out by the authors, impractical to compute, especially for large-size problems. We propose here a computable expression for the resulting condition number (exact formula and upper bound) using data that could be already available from the TLS solution process, namely, byproducts of the singular value decompositions (SVD) of A and $[A, b]$. We also make use of the adjoint operator which enables us to work on a space of lower dimension and which, to our knowledge, has never been used to derive TLS condition numbers. These adjoint techniques allow us to propose a practical algorithm based on the power method for computing the TLS condition number.

2. Definitions and notations.

2.1. The TLS problem. Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ with $m > n$. Following [17], we consider the two SVDs of A and $[A, b] : A = U'\Sigma'V'^T$ and $[A, b] = U\Sigma V^T$. We also set $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_{n+1})$, $\Sigma' = \mathrm{diag}(\sigma'_1, \ldots, \sigma'_n)$, where the singular values are in nonincreasing order, and we define $\lambda_i = \sigma_i^2$ and $\lambda_i' = \sigma_i'^2$. From [5, p. 178], we have the interlacing property

$$(2.1) \qquad \sigma_1 \geq \sigma'_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq \sigma'_n \geq \sigma_{n+1}.$$

We consider the TLS problem expressed in (1.1), and we assume in this text that the *genericity* condition $\sigma_n' > \sigma_{n+1}$ holds (for more information about the "nongeneric" problem, see, e.g., [17] and [22]). From [17, Theorems 2.6 and 2.7], it follows that the TLS solution x exists, is unique, and satisfies

$$(2.2) \qquad x = (A^T A - \lambda_{n+1} I_n)^{-1} A^T b.$$

In addition, $\begin{bmatrix} x \\ -1 \end{bmatrix}$ is an eigenvector of $[A, b]^T [A, b]$ associated with the simple eigenvalue λ_{n+1}; i.e., $\sigma'_n > \sigma_{n+1}$ guarantees that λ_{n+1} is not a semisimple eigenvalue of $[A, b]^T [A, b]$. As for linear LS problems, we define the TLS residual $r = b - Ax$, which enables us to write

$$(2.3) \qquad \lambda_{n+1} = \frac{1}{1 + x^T x} [x^T, \quad -1] \begin{bmatrix} A^T A & A^T b \\ b^T A & b^T b \end{bmatrix} \begin{bmatrix} x \\ -1 \end{bmatrix} = \frac{r^T r}{1 + x^T x}.$$

As mentioned in [17, p. 35], the TLS solution is obtained by scaling the last right singular vector v_{n+1} of $[A, b]$ until its last component is -1 and, if $v_{i,n+1}$ denotes the ith component of v_{n+1}, we have

$$(2.4) \qquad x = -\frac{1}{v_{n+1,n+1}} [v_{1,n+1}, \ldots, v_{n,n+1}]^T.$$

The TLS method involves an SVD computation, and the computational cost is higher than that of a classical LS problem (about $2mn^2 + 12n^3$ as mentioned in [13, p. 598], to be compared with the approximately $2mn^2$ flops required for LS solved via Householder QR factorization). However, there exist faster methods referred to as "partial SVD" (PSVD) that calculate only the last right singular vector or a basis of the right singular subspace associated with the smallest singular values of $[A, b]$ (see [17, p. 97]).

2.2. Condition number of the TLS problem. To measure the perturbations of the data A and b, we consider the product norm defined on $\mathbb{R}^{m \times n} \times \mathbb{R}^m$ by $\|(A, b)\|_F = \sqrt{\|A\|_F^2 + \|b\|_2^2}$, and we take the Euclidean norm $\|x\|_2$ for the solution space $\mathbb{R}^n$. In the following, the $n \times n$ identity matrix is denoted by I_n.

Let L be a given $n \times k$ matrix with $k \leq n$. We suppose here that L is not perturbed numerically, and we consider the mapping

$$g : \mathbb{R}^{m \times n} \times \mathbb{R}^m \to \mathbb{R}^k,$$
$$(A, b) \mapsto g(A, b) = L^T x = L^T (A^T A - \lambda_{n+1} I_n)^{-1} A^T b.$$

Since λ_{n+1} is simple, g is a Fréchet-differentiable function of A and b, and the genericity assumption ensures that the matrix $(A^T A - \lambda_{n+1} I_n)^{-1}$ is also Fréchet-differentiable in a

neighborhood of (A, b). As a result, g is Fréchet-differentiable in a neighborhood of (A, b).

The approach that we follow here is based on the work by [11] and [23], where the mathematical difficulty of a problem is measured by the norm of the Fréchet derivative of the problem solution expressed as a function of its data. This measure is an attainable bound in the limit as $(\Delta A, \Delta b) \to 0$, and it may therefore be approximate depending on the size of the perturbations. In general, the larger the ill-condition of the problem, the smaller the perturbations should be for this measure to provide a good bound, or guide, to the possible solution change. This will be further illustrated in a numerical example (section 4.2.1).

Using the definition given in [11] and [23], we can express the condition number of $L^T x$, a linear function of the TLS solution, as

$$(2.5) \qquad K(L, A, b) = \max_{(\Delta A, \Delta b) \neq 0} \frac{\|g'(A, b).(\Delta A, \Delta b)\|_2}{\|(\Delta A, \Delta b)\|_F}.$$

Note that the normwise condition number computed in [26, Theorem 3.1] corresponds to the same definition as formula (2.5).

$K(L, A, b)$ is sometimes called the *absolute* condition number of $L^T x$ as opposed to the *relative* condition number of $L^T x$ which is defined, when $L^T x$ is nonzero, by

$$(2.6) \qquad K^{(\mathrm{rel})}(L, A, b) = K(L, A, b)\|(A, b)\|_F / \|L^T x\|_2.$$

Using *relative* condition numbers can be useful when we are interested in *relative* forward and backward errors.

In what follows, the quantity $K(L, A, b)$ will be simply referred to as the TLS condition number, even though the proper conditioning of the TLS solution corresponds to the special case when L is the identity matrix. In the expression $g'(A, b).(\Delta A, \Delta b)$, the "." operator denotes that we apply the linear function $g'(A, b)$ to the variable $(\Delta A, \Delta b)$. We will use this notation throughout this paper to designate the image of a vector or a matrix by a linear function.

Remark 1. The case where $g(A, b) = h(x)$, with h being a differentiable nonlinear function mapping $\mathbb{R}^n$ to $\mathbb{R}^k$, is also covered because we have

$$g'(A, b).(\Delta A, \Delta b) = h'(x).(x'(A, b).(\Delta A, \Delta b)),$$

and L^T would correspond to the Jacobian matrix $h'(x)$. The nonlinear function h can be, for instance, the Euclidean norm of part of the solution (e.g., in the computation of Fourier coefficients when we are interested in the quantity of signal in a given frequency band).

3. Explicit formula for the TLS condition number.

3.1. Fréchet derivative. In this section, we compute the Fréchet derivative of g under the genericity assumption, which enables us to obtain an explicit formula for the TLS condition number in Proposition 2.

PROPOSITION 1. *Under the genericity assumption, g is Fréchet-differentiable in a neighborhood of (A, b). Setting $B_\lambda = A^T A - \lambda_{n+1} I_n$, the Fréchet derivative of g at (A, b) is expressed by*

$$g'(A, b) : \mathbb{R}^{m \times n} \times \mathbb{R}^m \to \mathbb{R}^k,$$

$$(\Delta A, \Delta b) \mapsto L^T B_\lambda^{-1} \left(A^T + \frac{2xr^T}{1 + x^T x} \right) (\Delta b - \Delta A x)$$

(3.1)
$$+ L^T B_\lambda^{-1} \Delta A^T r.$$

Proof. The result is obtained from the chain rule. Since λ_{n+1}, expressed in (2.3), is a simple eigenvalue of the symmetric matric $[A, b]^T [A, b]$ with corresponding unit eigenvector $\frac{1}{\sqrt{1+x^T x}} [x^T \quad -1]^T$, λ_{n+1} is differentiable in a neighborhood of (A, b), and then we know from [24, p. 45] that the derivative of λ_{n+1} can be expressed as a function of the first order perturbation of the matrix $[A, b]^T [A, b]$ and of the normalized eigenvector $\frac{1}{\sqrt{1+x^T x}} [x^T \quad -1]^T$, where x is the TLS solution of the nonperturbed problem:

$$\lambda'_{n+1}(A, b).(\Delta A, \Delta b) = \frac{1}{1 + x^T x} [x^T \quad -1] \begin{bmatrix} \Delta A^T A + A^T \Delta A & \Delta A^T b + A^T \Delta b \\ b^T \Delta A + \Delta b^T A & \Delta b^T b + b^T \Delta b \end{bmatrix} \begin{bmatrix} x \\ -1 \end{bmatrix}$$

$$= \frac{2}{1 + x^T x} (x^T \Delta A^T A x - x^T \Delta A^T b - x^T A^T \Delta b + b^T \Delta b)$$

$$= \frac{2}{1 + x^T x} (-x^T \Delta A^T r + (b^T - x^T A^T) \Delta b),$$

$$= \frac{2}{1 + x^T x} (-r^T \Delta A x + r^T \Delta b),$$

yielding

(3.2)
$$\lambda'_{n+1}(A, b).(\Delta A, \Delta b) = \frac{2r^T (\Delta b - \Delta A x)}{1 + x^T x}.$$

Since, for a nonsingular matrix M, $\partial(M^{-1})/\partial\tau = -M^{-1}(\partial M/\partial\tau)M^{-1}$, we obtain

$$(B_\lambda^{-1})'(A, b).(\Delta A, \Delta b) = -B_\lambda^{-1}(\Delta A^T A + A^T \Delta A - \lambda'_{n+1}(A, b).(\Delta A, \Delta b)I_n)B_\lambda^{-1}$$

$$= -B_\lambda^{-1} \left(\Delta A^T A + A^T \Delta A - \frac{2r^T (\Delta b - \Delta A x)}{1 + x^T x} I_n \right) B_\lambda^{-1}.$$

The chain rule now applied to $g(A, b)$ leads to

$$g'(A, b).(\Delta A, \Delta b) = -L^T B_\lambda^{-1}(\Delta A^T A + A^T \Delta A - \lambda'_{n+1}(A, b).(\Delta A, \Delta b)I_n)B_\lambda^{-1} A^T b$$

$$+ L^T B_\lambda^{-1}(\Delta A^T b + A^T \Delta b)$$

$$= -L^T B_\lambda^{-1}(\Delta A^T A + A^T \Delta A - \lambda'_{n+1}(A, b).(\Delta A, \Delta b)I_n)x$$

$$+ L^T B_\lambda^{-1}(\Delta A^T b + A^T \Delta b)$$

$$= L^T B_\lambda^{-1} \left(A^T + \frac{2xr^T}{1 + x^T x} \right) (\Delta b - \Delta A x) + L^T B_\lambda^{-1} \Delta A^T r,$$

which gives the result.   □

We now introduce the vec operation that stacks all the columns of a matrix into a long vector: for $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$, $\mathrm{vec}(A) = [a_1^T, \ldots, a_n^T]^T \in \mathbb{R}^{mn \times 1}$. Let $P \in \mathbb{R}^{mn \times mn}$ denote the permutation matrix that represents the matrix transpose by $\mathrm{vec}(B^T) = P\mathrm{vec}(B)$. We remind the reader that $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$, where $\otimes$ denotes the Kronecker product of two matrices [14, p. 21].

Let us now obtain the matrix $\mathcal{M}_{g'}$ representing $g'(A, b)$, i.e., such that

$$(3.3) \qquad g'(A, b).(\Delta A, \Delta b) = \mathcal{M}_{g'} \begin{bmatrix} \mathrm{vec}(\Delta A) \\ \Delta b \end{bmatrix}.$$

Since $g'(A, b).(\Delta A, \Delta b) \in \mathbb{R}^k$, we have $g'(A, b).(\Delta A, \Delta b) = \mathrm{vec}(g'(A, b).(\Delta A, \Delta b))$, and setting in addition $D_\lambda = L^T B_\lambda^{-1}(A^T + \frac{2xr^T}{1+x^Tx}) \in \mathbb{R}^{k \times m}$, we obtain from (3.1)

$$\begin{aligned} g'(A, b).(\Delta A, \Delta b) &= \mathrm{vec}(D_\lambda(\Delta b - \Delta Ax) + L^T B_\lambda^{-1}\Delta A^T r) \\ &= (-x^T \otimes D_\lambda)\mathrm{vec}(\Delta A) + (r^T \otimes (L^T B_\lambda^{-1}))\mathrm{vec}(\Delta A^T) + D_\lambda \Delta b \\ &= [-x^T \otimes D_\lambda + (r^T \otimes (L^T B_\lambda^{-1}))P, D_\lambda] \begin{bmatrix} \mathrm{vec}(\Delta A) \\ \Delta b \end{bmatrix}. \end{aligned}$$

Then we get

$$\mathcal{M}_{g'} = [-x^T \otimes D_\lambda + (r^T \otimes (L^T B_\lambda^{-1}))P, D_\lambda] \in \mathbb{R}^{k \times (nm+m)}.$$

But we have $\|(\Delta A, \Delta b)\|_F = \|\begin{bmatrix} \mathrm{vec}(\Delta A) \\ \Delta b \end{bmatrix}\|_2$ and then, from Proposition 1 and using the definition of $K(L, A, b)$ given in (2.5), we get the following proposition that expresses the TLS condition number in terms of the norm of a matrix.

PROPOSITION 2. *The condition number of $g(A, b)$ is given by*

$$K(L, A, b) = \|\mathcal{M}_{g'}\|_2,$$

where

$$\mathcal{M}_{g'} = [-x^T \otimes D_\lambda + (r^T \otimes (L^T B_\lambda^{-1}))P, D_\lambda] \in \mathbb{R}^{k \times (nm+m)}.$$

3.2. Adjoint operator and algorithm. Computing $K(L, A, b)$ reduces to computing the spectral norm of the $k \times (nm + m)$ matrix $\mathcal{M}_{g'}$. For large values of n or m, it is not possible to build explicitly the generally dense matrix $\mathcal{M}_{g'}$. Iterative techniques based on the power method [16, p. 289] or on the Lanczos method [13] are better suited. These algorithms involve, however, the computation of the product of $\mathcal{M}_{g'}^T$ by a vector $y \in \mathbb{R}^k$. We describe now how to perform this operation.

Using successively the fact that $B_\lambda^{-T} = B_\lambda^{-1}$, $(A \otimes B)^T = A^T \otimes B^T$, $\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X)$, and $P^T = P^{-1}$, we have

$$\mathcal{M}_{g'}^T y = \begin{bmatrix} -x \otimes D_\lambda^T + P^T(r \otimes (B_\lambda^{-T}L)) \\ D_\lambda^T \end{bmatrix} y$$

$$= \begin{bmatrix} -(x \otimes D_\lambda^T)\mathrm{vec}(y) + P^T(r \otimes (B_\lambda^{-1}L))\mathrm{vec}(y) \\ D_\lambda^T y \end{bmatrix}$$

$$= \begin{bmatrix} P^{-1}(P\mathrm{vec}(-D_\lambda^T yx^T) + \mathrm{vec}(B_\lambda^{-1}Lyr^T)) \\ D_\lambda^T y \end{bmatrix}$$

$$= \begin{bmatrix} P^{-1}(\mathrm{vec}((-D_\lambda^T yx^T)^T) + \mathrm{vec}(B_\lambda^{-1}Lyr^T)) \\ D_\lambda^T y \end{bmatrix}$$

$$= \begin{bmatrix} P^{-1}\mathrm{vec}(-xy^T D_\lambda + B_\lambda^{-1}Lyr^T) \\ D_\lambda^T y \end{bmatrix},$$

and since for any matrix B we have $P^{-1}\mathrm{vec}(B) = \mathrm{vec}(B^T)$, we get

$$(3.4) \qquad \mathcal{M}_{g'}^T y = \begin{bmatrix} \mathrm{vec}(-D_\lambda^T yx^T + ry^T L^T B_\lambda^{-1}) \\ D_\lambda^T y \end{bmatrix}.$$

This leads us to the following proposition.

PROPOSITION 3. *The adjoint operator of $g'(A, b)$, using the scalar products* $\mathrm{trace}(A_1^T A_2) + b_1^T b_2$ *and* $y_1^T y_2$, *respectively, on $\mathbb{R}^{m \times n} \times \mathbb{R}^m$ and $\mathbb{R}^k$, is*

$$g'^*(A, b) : \mathbb{R}^k \to \mathbb{R}^{m \times n} \times \mathbb{R}^m,$$

$$(3.5) \qquad y \mapsto (-D_\lambda^T yx^T + ry^T L^T B_\lambda^{-1}, D_\lambda^T y).$$

In addition, if $k = 1$, we have

$$(3.6) \qquad K(L, A, b) = \sqrt{\| -D_\lambda^T x^T + rL^T B_\lambda^{-1} \|_F^2 + \|D_\lambda\|_2^2}.$$

Proof. Let us denote by $\langle (A_1, b_1), (A_2, b_2) \rangle$ the scalar product $\mathrm{trace}(A_1^T A_2) + b_1^T b_2$ on $\mathbb{R}^{m \times n} \times \mathbb{R}^m$. We have, for any $y \in \mathbb{R}^k$, using (3.3), then (3.4),

$$y^T(g'(A, b).(\Delta A, \Delta b)) = y^T \mathcal{M}_{g'} \begin{bmatrix} \mathrm{vec}(\Delta A) \\ \Delta b \end{bmatrix}$$

$$= (\mathcal{M}_{g'}^T y)^T \begin{bmatrix} \mathrm{vec}(\Delta A) \\ \Delta b \end{bmatrix}$$

$$= \mathrm{vec}(-D_\lambda^T yx^T + ry^T L^T B_\lambda^{-1})^T \mathrm{vec}(\Delta A) + (D_\lambda^T y)^T \Delta b.$$

Using now the fact that, for matrices A_1 and A_2 of identical sizes, $\mathrm{vec}(A_1)^T \mathrm{vec}(A_2) = \mathrm{trace}(A_1^T A_2)$, we get

$$
\begin{aligned}
y^T(g'(A, b).(\Delta A, \Delta b)) &= \mathrm{trace}((-D_\lambda^T y x^T + r y^T L^T B_\lambda^{-1})^T \Delta A) + (D_\lambda^T y)^T \Delta b \\
&= \langle (-D_\lambda^T y x^T + r y^T L^T B_\lambda^{-1}, D_\lambda^T y), (\Delta A, \Delta b) \rangle \\
&= \langle g'^*(A, b).y, (\Delta A, \Delta b) \rangle,
\end{aligned}
$$

which concludes the first part of the proof.

For the second part, we use (3.4) to give

$$
K(L, A, b) = \|\mathcal{M}_{g'}\|_2 = \|\mathcal{M}_{g'}^T\|_2 = \max_{y \neq 0} \frac{\left\| \begin{bmatrix} \mathrm{vec}(-D_\lambda^T y x^T + r y^T L^T B_\lambda^{-1}) \\ D_\lambda^T y \end{bmatrix} \right\|_2}{\|y\|_2}.
$$

Since $k = 1$, we have $y \in \mathbb{R}$ and $K(L, A, b) = \left\| \begin{bmatrix} \mathrm{vec}(-D_\lambda^T x^T + r L^T B_\lambda^{-1}) \\ D_\lambda^T \end{bmatrix} \right\|_2$, and the result follows from the relation $\mathrm{vec}(A_1)^T \mathrm{vec}(A_1) = \mathrm{trace}\, A_1^T A_1 = \|A_1\|_F^2$. $\square$

Remark 2. The special case $k = 1$ covers the situation where we compute the conditioning of the ith solution component. In that case L is the ith canonical vector of $\mathbb{R}^n$ and, in (3.6), $L^T B_\lambda^{-1}$ is the ith row of B_λ^{-1} and D_λ is the ith row of $B_\lambda^{-1}(A^T + \frac{2xr^T}{1+x^T x})$.

Using (3.1) and (3.5), we can now write in Algorithm 1 the iteration of the power method [16, p. 289] to compute the TLS condition number $K(L, A, b)$. In this algorithm we assume x and λ_{n+1} are available, and we iterate (A_p, b_p) to approach the optimal $(\Delta A, \Delta b)$ that realizes (2.5).

ALGORITHM 1. CONDITION NUMBER OF TLS PROBLEM.

Select initial vector $y \in \mathbb{R}^k$.

for p = 1, 2, ...

$\quad (A_p, b_p) = (-D_\lambda^T y x^T + r y^T L^T B_\lambda^{-1}, D_\lambda^T y)$ {(3.5)}

$\quad \nu = \|(A_p, b_p)\|_F$

$\quad (A_p, b_p) \leftarrow (\frac{1}{\nu} \cdot A_p, \frac{1}{\nu} \cdot b_p)$

$\quad y = L^T B_\lambda^{-1}(A^T + \frac{2xr^T}{1+x^T x})(b_p - A_p x) + L^T B_\lambda^{-1} A_p^T r$ {(3.1)}

end

$K(L, A, b) = \sqrt{\nu}$

The quantity ν computed by Algorithm 1 is the largest eigenvalue of $\mathcal{M}_{g'} \mathcal{M}_{g'}^T$. Since $K(L, A, b) = \|\mathcal{M}_{g'}\|_2$, the condition number $K(L, A, b)$ is also the largest singular value of $\mathcal{M}_{g'}$, i.e., $\sqrt{\nu}$. As mentioned in [13, p. 331], the algorithm will converge if the initial y has a component in the direction of the corresponding dominant eigenvector of $\mathcal{M}_{g'} \mathcal{M}_{g'}^T$. When there is an estimate of this dominant eigenvector, the initial y can be set to this estimate, but in many implementations y is initialized as a random vector. The algorithm is terminated by a "sufficiently" large number of iterations or by evaluating the difference between two successive values of ν and comparing it to a tolerance given by the user.

3.3. Closed formula. Using the adjoint formulas obtained in section 3.2, we now get a closed formula for the TLS conditioning.

THEOREM 1. *We consider the TLS problem and assume that the genericity assumption holds. Setting $B_\lambda = A^T A - \lambda_{n+1} I_n$, then from Proposition 2 the condition number of $L^T x$, a linear function of the TLS solution, is expressed by*

$$K(L, A, b) = \| C \|_2^{\frac{1}{2}},$$

where C is the $k \times k$ symmetric matrix

$$C = \mathcal{M}_{g'} \mathcal{M}_{g'}^T = (1 + \|x\|_2^2) L^T B_\lambda^{-1} \left(A^T A + \lambda_{n+1} \left(I_n - \frac{2xx^T}{1 + \|x\|_2^2} \right) \right) B_\lambda^{-1} L.$$

Proof. From Proposition 2 we have $C = \mathcal{M}_{g'} \mathcal{M}_{g'}^T$ with

$$\mathcal{M}_{g'} = [-x^T \otimes D_\lambda + (r^T \otimes (L^T B_\lambda^{-1})) P, \quad D_\lambda].$$

Then, using $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ and $P(x \otimes D_\lambda^T) = D_\lambda^T \otimes x$, we get

$$\begin{aligned} C &= [-x^T \otimes D_\lambda + (r^T \otimes (L^T B_\lambda^{-1})) P][-x \otimes D_\lambda^T + P^T (r \otimes (B_\lambda^{-1} L)) P] + D_\lambda D_\lambda{}^T \\ &= (x^T x) \otimes (D_\lambda D_\lambda^T) + (r^T \otimes (L^T B_\lambda^{-1}))(r \otimes (B_\lambda^{-1} L)) \\ &\quad - 2(r^T \otimes (L^T B_\lambda^{-1})) P(x \otimes D_\lambda{}^T) + D_\lambda D_\lambda{}^T \\ &= (x^T x) \otimes (D_\lambda D_\lambda^T) + D_\lambda D_\lambda^T + (r^T r) \otimes (L^T B_\lambda^{-2} L) - 2(r^T \otimes (L^T B_\lambda^{-1}))(D_\lambda^T \otimes x) \\ &= (1 + x^T x) D_\lambda D_\lambda^T + \|r\|_2^2 L^T B_\lambda^{-2} L - 2(r^T D_\lambda^T)(L^T B_\lambda^{-1} x), \end{aligned}$$

and using the fact that, for two vectors $y^T \otimes z = zy^T$, we obtain

$$(3.7) \qquad C = (1 + x^T x) D_\lambda D_\lambda^T + \|r\|_2^2 L^T B_\lambda^{-2} L - 2 L^T B_\lambda^{-1} x r^T D_\lambda^T.$$

Replacing D_λ by $L^T B_\lambda^{-1}(A^T + \frac{2xr^T}{1 + x^T x})$, (3.7) simplifies to

$$(3.8) \qquad C = L^T B_\lambda^{-1}((1 + x^T x) A^T A + \|r\|_2^2 I_n + 2 A^T r x^T) B_\lambda^{-1} L.$$

But $A^T r x^T = A^T(b - Ax)x^T = A^T b x^T - A^T A x x^T$ and, since from (2.2) we have $A^T b = B_\lambda x$, we get $A^T r x^T = B_\lambda x x^T - A^T A x x^T = (A^T A - \lambda_{n+1} I_n) x x^T - A^T A x x^T = -\lambda_{n+1} x x^T$. From (2.3) we also have $\|r\|_2^2 = \lambda_{n+1}(1 + x^T x)$, and thus (3.8) becomes

$$\begin{aligned} C &= L^T B_\lambda^{-1}((1 + x^T x) A^T A + \lambda_{n+1}(1 + x^T x) I_n - 2\lambda_{n+1} x x^T) B_\lambda^{-1} L \\ &= (1 + \|x\|_2^2) L^T B_\lambda^{-1} \left(A^T A + \lambda_{n+1} \left(I_n - \frac{2xx^T}{1 + \|x\|_2^2} \right) \right) B_\lambda^{-1} L. \qquad \square \end{aligned}$$

4. TLS condition number and SVD.

4.1. Closed formula and upper bound. Computing $K(L, A, b)$ using Theorem 1 requires the explicit formation of the normal equations matrix $A^T A$ which is a source of rounding errors and also generates an extra computational cost of about mn^2 flops. In practice the TLS solution is obtained by (2.4) and involves an SVD computation. In the following theorem, we propose a formula for $K(L, A, b)$ that can be computed with quantities that may be already available from the solution process. In the following $0_{n,1}$ (resp., $0_{1,n}$) denotes the zero column (resp., row) vector of length n.

THEOREM 2. *Let V and V' be the matrices whose columns are the right singular vectors of, respectively, $[A, b]$ and A associated with the singular values $(\sigma_1, \ldots, \sigma_{n+1})$ and $(\sigma'_1, \ldots, \sigma'_n)$. Then the condition number of $L^T x$, a linear function of the TLS solution, is expressed by*

$$K(L, A, b) = (1 + \|x\|_2^2)^{\frac{1}{2}} \|L^T V' D'[V'^T, 0_{n,1}] V[D, 0_{n,1}]^T\|_2, \quad where$$

$$D' = \mathrm{diag}((\sigma_1'^2 - \sigma_{n+1}^2)^{-1}, \ldots, (\sigma_n'^2 - \sigma_{n+1}^2)^{-1}) \quad and$$

$$D = \mathrm{diag}((\sigma_1^2 + \sigma_{n+1}^2)^{\frac{1}{2}}, \ldots, (\sigma_n^2 + \sigma_{n+1}^2)^{\frac{1}{2}}).$$

When L is the identity matrix, then the condition number reduces to

$$K(L, A, b) = (1 + \|x\|_2^2)^{\frac{1}{2}} \|D'[V'^T, 0_{n,1}] V[D, 0_{n,1}]^T\|_2.$$

Proof. From $[A, b] = U\Sigma V^T$, we have $[A, b]^T[A, b] = V\Sigma^2 V^T = \sum_{i=1}^{n+1} \sigma_i^2 v_i v_i^T$ and

$$[A, b]^T[A, b] + \lambda_{n+1} I_{n+1} = \sum_{i=1}^{n+1} \sigma_i^2 v_i v_i^T + \lambda_{n+1} \sum_{i=1}^{n+1} v_i v_i^T$$

$$= \sum_{i=1}^{n+1} (\sigma_i^2 + \lambda_{n+1}) v_i v_i^T$$

$$= \sum_{i=1}^{n} (\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T + 2\lambda_{n+1} v_{n+1} v_{n+1}^T,$$

leading to

$$(4.1) \qquad [A, b]^T[A, b] + \lambda_{n+1} I_{n+1} - 2\lambda_{n+1} v_{n+1} v_{n+1}^T = \sum_{i=1}^{n} (\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T.$$

From (2.4) we have $v_{n+1} = -v_{n+1,n+1}\begin{bmatrix} x \\ -1 \end{bmatrix}$ and, since v_{n+1} is a unit vector, $v_{n+1,n+1}^2 = \frac{1}{1+\|x\|_2^2}$. Then (4.1) can be expressed in matrix notation as

$$\begin{bmatrix} A^T A & A^T b \\ b^T A & b^T b \end{bmatrix} + \lambda_{n+1} \begin{bmatrix} I_n & 0_{n,1} \\ 0_{1,n} & 1 \end{bmatrix} - \frac{2\lambda_{n+1}}{1+\|x\|_2^2} \begin{bmatrix} xx^T & -x \\ -x^T & 1 \end{bmatrix} = \sum_{i=1}^{n} (\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T.$$

(4.2)

The quantity $A^T A + \lambda_{n+1}(I_n - \frac{2xx^T}{1+\|x\|_2^2})$ corresponds to the left-hand side of (4.2) in which the last row and the last column have been removed. Thus it can also be written as

$$A^T A + \lambda_{n+1}\left(I_n - \frac{2xx^T}{1+\|x\|_2^2}\right) = [I_n, 0_{n,1}]\left(\sum_{i=1}^{n}(\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T\right)\begin{bmatrix} I_n \\ 0_{1,n} \end{bmatrix},$$

and the matrix C from Theorem 1 can be expressed as

$$(4.3) \qquad C = (1 + \|x\|_2^2) L^T[B_\lambda^{-1}, 0_{n,1}]\left(\sum_{i=1}^{n}(\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T\right)\begin{bmatrix} B_\lambda^{-1} \\ 0_{1,n} \end{bmatrix} L.$$

Moreover, from $A = U'\Sigma' V'^T$, we have $A^T A = V'\Sigma'^2 V'^T = \sum_{i=1}^{n} \sigma_i'^2 v_i' v_i'^T$ and

$$B_\lambda = A^T A - \lambda_{n+1} I_n$$

$$= \sum_{i=1}^{n} \sigma_i'^2 v_i' v_i'^T - \sigma_{n+1}^2 \sum_{i=1}^{n} v_i' v_i'^T$$

$$= \sum_{i=1}^{n} (\sigma_i'^2 - \sigma_{n+1}^2) v_i' v_i'^T$$

$$= V' D'^{-1} V'^T.$$

Hence $B_\lambda^{-1} = V'^{-T} D' V'^{-1} = V' D' V'^T$ and $[B_\lambda^{-1}, 0_{n,1}] = V' D' [V'^T, 0_{n,1}]$. We also have $\sum_{i=1}^{n} (\sigma_i^2 + \sigma_{n+1}^2) v_i v_i^T = V \begin{bmatrix} D \\ 0_{1,n} \end{bmatrix} [D, 0_{n,1}] V^T$. Then, by substituting these in (4.3), we obtain $C = (1 + \|x\|_2^2) \tilde{V} \tilde{V}^T$ with $\tilde{V} = L^T V' D' [V'^T, 0_{n,1}] V [D, 0_{n,1}]^T$. As a result, using Theorem 1,

$$K(L, A, b)^2 = \|C\|_2 = (1 + \|x\|_2^2) \|\tilde{V} \tilde{V}^T\|_2 = (1 + \|x\|_2^2) \|\tilde{V}\|_2^2.$$

When $L = I_n$, we use the fact that V' is an orthogonal matrix and can be removed from the expression of $\|\tilde{V}\|_2^2$.   □

In many applications, an upper bound would be sufficient to give an estimate of the conditioning of the TLS solution. The following corollary gives an upper bound for $K(L, A, b)$.

COROLLARY 1. *The condition number of $L^T x$, a linear function of the TLS solution, is bounded by*

$$\bar{K}(L, A, b) = (1 + \|x\|_2^2)^{\frac{1}{2}} \|L\|_2 \frac{(\sigma_1^2 + \sigma_{n+1}^2)^{\frac{1}{2}}}{(\sigma_n'^2 - \sigma_{n+1}^2)}.$$

Proof. This result comes from the inequality $\|AB\|_2 \le \|A\|_2 \|B\|_2$, followed by $\|D'\|_2 = \max_i (\sigma_i'^2 - \sigma_{n+1}^2)^{-1} = (\sigma_n'^2 - \sigma_{n+1}^2)^{-1}$ and $\|D\|_2^2 = \max_i (\sigma_i^2 + \sigma_{n+1}^2) = (\sigma_1^2 + \sigma_{n+1}^2)$.   □

Note that the three quantities used for computing $\bar{K}(L, A, b)$ (i.e., σ_n', σ_{n+1}, and σ_1) can be obtained from the orthogonal upper bidiagonalization of $[b, A]$ (see [22]).

4.2. Numerical examples. In the following examples we study the condition number of x; i.e., L is here the identity matrix I_n. Then, to simplify the notation, we remove the variable L from the expressions and the condition number of x will be denoted by $K(A, b)$ and its upper bound by $\bar{K}(A, b)$. All the experiments were performed with MATLAB 7.6.0 using a machine precision $2.22 \cdot 10^{-16}$.

In Example 1, we compare $K(A, b)$ with bounds coming from Corollary 1 and from the literature. In Example 2, for a well-conditioned problem we compare forward errors with bounds on them predicted by $K(A, b)$ and by upper bounds on $K(A, b)$. The aim of Example 3 is to show the limitations of the first order approach used in the condition number definition (2.5) when seeking forward error bounds.

4.2.1. Example 1. In this first example we consider the TLS problem $Ax \approx b$, where $[A, b]$ is defined by

$$[A, b] = Y \begin{pmatrix} D \\ 0 \end{pmatrix} Z^T \in \mathbb{R}^{m \times (n+1)}, \quad Y = I_m - 2yy^T, \quad Z = I_{n+1} - 2zz^T,$$

where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^{n+1}$ are random unit vectors and $D = \mathrm{diag}(n, n-1, \ldots, 1, 1 - e_p)$ for a given parameter e_p. The quantity $\sigma'_n - \sigma_{n+1}$ measures the distance of our problem to nongenericity and, due to (2.1), we have in exact arithmetic

$$\sigma'_n - \sigma_{n+1} \leq \sigma_n - \sigma_{n+1} = e_p.$$

Then by varying e_p, we can generate different TLS problems, and by considering small values of e_p, it is possible to study the behavior of the TLS condition number in the context of close-to-nongeneric problems. The TLS solution x is computed using an SVD of $[A, b]$ and (2.4).

We consider the values $m = 100$, $n = 20$, and we compare in Table 4.1 the exact condition number $K(A, b)$ given in Theorem 2, the upper bound $\bar{K}(A, b)$ given in Corollary 1, and the upper bound obtained from [17, p. 212] and expressed by

$$\kappa(A, b) = \frac{9\sigma_1 \|x\|_2}{\sigma_n - \sigma_{n+1}} \left(1 + \frac{\|b\|_2}{\sigma'_n - \sigma_{n+1}}\right) \frac{1}{\|b\|_2 - \sigma_{n+1}}.$$

We also report the condition number computed by Algorithm 1, denoted by $K_p(A, b)$, and the corresponding number of power iterations (the algorithm terminates when the difference between two successive values is lower than 10^{-8}). When $\sigma'_n - \sigma_{n+1}$ decreases, the TLS problem becomes more poorly conditioned, and there is a factor $\mathcal{O}(10)$ between the exact condition number $K(A, b)$ and its upper bound $\bar{K}(A, b)$. We also observe that the estimate $\bar{K}(A, b)$ can be many orders of magnitude better than $\kappa(A, b)$ and that, for small values of $\sigma'_n - \sigma_{n+1}$, $\kappa(A, b)$ is much less reliable. $K_p(A, b)$ is always equal or very close to $K(A, b)$.

4.2.2. Example 2. Let us now consider the following example from [17, p. 42], also used in [26], where

$$A = \begin{pmatrix} m-1 & -1 & \cdots & -1 \\ -1 & m-1 & \cdots & -1 \\ \vdots & & & \\ -1 & -1 & \cdots & m-1 \\ -1 & -1 & \cdots & -1 \\ -1 & -1 & \cdots & -1 \end{pmatrix} \in \mathbb{R}^{m \times (m-2)}, \qquad b = \begin{pmatrix} -1 \\ -1 \\ \vdots \\ -1 \\ m-1 \\ -1 \end{pmatrix} \in \mathbb{R}^m.$$

The exact solution of the TLS problem $Ax \approx b$ can be computed analytically [17, p. 42] and is equal to $x = -(1, \ldots, 1)^T$. We consider a random perturbation $(\Delta A, \Delta b)$ of small

TABLE 4.1
TLS conditioning for several values of $\sigma'_n - \sigma_{n+1}$ ($m = 100$, $n = 20$).

$\sigma'_n - \sigma_{n+1}$	$K(A, b)$	$\bar{K}(A, b)$	$\kappa(A, b)$	$K_p(A, b)$	#iter
$9.99976032 \cdot 10^{-1}$	$1.18 \cdot 10^0$	$2.36 \cdot 10^1$	$1.29 \cdot 10^2$	$1.18 \cdot 10^0$	11
$9.99952397 \cdot 10^{-5}$	$8.36 \cdot 10^3$	$1.18 \cdot 10^5$	$1.31 \cdot 10^{10}$	$8.36 \cdot 10^3$	6
$9.99952365 \cdot 10^{-9}$	$8.36 \cdot 10^7$	$1.18 \cdot 10^9$	$1.31 \cdot 10^{18}$	$8.36 \cdot 10^7$	4
$9.99644811 \cdot 10^{-13}$	$8.36 \cdot 10^{11}$	$1.18 \cdot 10^{13}$	$1.31 \cdot 10^{26}$	$8.32 \cdot 10^{11}$	5

norm $\|(\Delta A, \Delta b)\|_F = 10^{-9}$, and we denote by $\tilde{x}$ the computed solution of the perturbed system $(A + \Delta A)x \approx b + \Delta b$.

In Table 4.2, we compare for several values of m the computed forward error $\|\tilde{x} - x\|_2$ with the forward error bounds that can be expected from the computation of $K(A, b)$ and its upper bounds $\bar{K}(A, b)$ and $\kappa(A, b)$. Note that this problem is very well-conditioned because the condition number $K(A, b)$ computed using Theorem 2 is close to 1 for each value of m. Since to first order the condition number corresponds to the worst-case error amplification, $K(A, b)\|(\Delta A, \Delta b)\|_F$ is, as observed in Table 4.2, always larger than the computed forward error (there is approximately a factor up to $\mathcal{O}(10^2)$ between those quantities). We also observe that, in this example, $\bar{K}(A, b)$ and $\kappa(A, b)$ produce forward error estimates that are of the same order of magnitude except in the case $m = 5$ for which $\kappa(A, b)$ is more pessimistic.

4.2.3. Example 3. As explained in section 2.2, the condition number, defined as the norm of the Fréchet derivative of the solution, is a first order term. In the following example, we show the limitation of this approach in providing good error bounds, depending on the conditioning of the problem and on the size of the perturbations.

We consider the same TLS problem $Ax \approx b$ as in Example 1 with the same dimensions of matrices and the same values for the parameter e_p, enabling us to vary the conditioning of the problem. For each value of e_p, we consider random relative perturbations $(\Delta A, \Delta b)$ such that $\frac{\|(\Delta A, \Delta b)\|_F}{\|(A, b)\|_F} = 10^{-q}$. Note that, for this problem, the exact solution $x = g(A, b)$ is known by construction and, with the notation of Example 1, is equal to $Z(1{:}n, n+1)/Z(n+1, n+1)$. Let $\tilde{x}$ be the computed solution.

For several values of $(e_p, 10^{-q})$, we report in Table 4.3 the condition number $K(A, b)$, the relative forward error, the relative error at first order, and the worst-case relative error estimate (that corresponds to the product of the relative condition number by the relative perturbation of data).

The rows of Table 4.3 are sorted by increasing condition numbers and increasing relative perturbations. When the problem is well-conditioned, we observe that the first order term $\|g'(A, b).(\Delta A, \Delta b)\|_2/\|x\|_2$ enables us to predict the relative forward error for all sizes of perturbations considered here. When the condition number increases, only small perturbations provide consistency between the forward error and the first order error. This indicates that, the larger the ill-condition of the problem, the less reliable the first order approach is when large perturbations are considered. We also notice that, for all experiments, the error estimate $K(A, b) \times 10^{-q}/\|x\|_2$ based on the condition number overestimates the first order error $\|g'(A, b).(\Delta A, \Delta b)\|_2/\|x\|_2$ with an order of magnitude $\mathcal{O}(10)$, which corresponds to the ratio between $\|g'(A, b)\| \times \|(\Delta A, \Delta b)\|_F$ and $\|g'(A, b).(\Delta A, \Delta b)\|_2$, where $\|g'(A, b)\|$ denotes the operator norm of the linear function $g'(A, b)$. In this particular example, we observe that, when

TABLE 4.2
Forward error and upper bounds for a perturbed TLS problem.

m	$\|\tilde{x} - x\|_2$	$K(A, b)\|(\Delta A, \Delta b)\|_F$	$\bar{K}(A, b)\|(\Delta A, \Delta b)\|_F$	$\kappa(A, b)\|(\Delta A, \Delta b)\|_F$
5	$1.50 \cdot 10^{-10}$	$1.72 \cdot 10^{-9}$	$3.45 \cdot 10^{-9}$	$1.16 \cdot 10^{-7}$
60	$1.59 \cdot 10^{-11}$	$1.53 \cdot 10^{-9}$	$1.18 \cdot 10^{-8}$	$4.52 \cdot 10^{-8}$
200	$8.09 \cdot 10^{-12}$	$1.53 \cdot 10^{-9}$	$2.16 \cdot 10^{-8}$	$3.93 \cdot 10^{-8}$

TABLE 4.3
First order estimation for TLS forward error for various conditioning and perturbation sizes.

e_p	q	$K(A,b)$	$\frac{\|\tilde{x}-x\|_2}{\|x\|_2}$	$\frac{\|g'(A,b).(\Delta A,\Delta b)\|_2}{\|x\|_2}$	$\frac{K(A,b)\times 10^{-q}}{\|x\|_2}$
$1 \cdot 10^0$	15	$1.18 \cdot 10^0$	$5.61 \cdot 10^{-15}$	$5.61 \cdot 10^{-15}$	$1.00 \cdot 10^{-13}$
	12	"	$4.82 \cdot 10^{-12}$	$4.82 \cdot 10^{-12}$	$1.00 \cdot 10^{-10}$
	9	"	$5.67 \cdot 10^{-9}$	$5.67 \cdot 10^{-9}$	$1.00 \cdot 10^{-7}$
	6	"	$4.50 \cdot 10^{-6}$	$4.50 \cdot 10^{-6}$	$1.00 \cdot 10^{-4}$
	3	"	$3.54 \cdot 10^{-3}$	$3.71 \cdot 10^{-3}$	$1.00 \cdot 10^{-1}$
$1 \cdot 10^{-4}$	15	$8.36 \cdot 10^3$	$2.29 \cdot 10^{-11}$	$2.22 \cdot 10^{-11}$	$7.10 \cdot 10^{-10}$
	12	"	$8.46 \cdot 10^{-9}$	$8.46 \cdot 10^{-9}$	$7.10 \cdot 10^{-7}$
	9	"	$1.99 \cdot 10^{-5}$	$1.99 \cdot 10^{-5}$	$7.10 \cdot 10^{-4}$
	6	"	$2.10 \cdot 10^{-2}$	$2.06 \cdot 10^{-2}$	$7.10 \cdot 10^{-1}$
	3	"	$8.95 \cdot 10^0$	$1.09 \cdot 10^1$	$7.10 \cdot 10^2$
$1 \cdot 10^{-8}$	15	$8.36 \cdot 10^7$	$2.11 \cdot 10^{-7}$	$2.22 \cdot 10^{-7}$	$7.10 \cdot 10^{-6}$
	12	"	$8.47 \cdot 10^{-5}$	$8.46 \cdot 10^{-5}$	$7.10 \cdot 10^{-3}$
	9	"	$2.31 \cdot 10^{-1}$	$1.99 \cdot 10^{-1}$	$7.10 \cdot 10^0$
	6	"	$4.39 \cdot 10^0$	$2.06 \cdot 10^2$	$7.10 \cdot 10^3$
	3	"	$9.27 \cdot 10^0$	$1.09 \cdot 10^5$	$7.10 \cdot 10^6$
$1 \cdot 10^{-12}$	15	$8.36 \cdot 10^{11}$	$3.61 \cdot 10^{-3}$	$2.21 \cdot 10^{-3}$	$7.11 \cdot 10^{-2}$
	12	"	$7.39 \cdot 10^0$	$8.42 \cdot 10^{-1}$	$7.11 \cdot 10^1$
	9	"	$3.62 \cdot 10^0$	$1.98 \cdot 10^3$	$7.11 \cdot 10^4$
	6	"	$4.40 \cdot 10^0$	$2.05 \cdot 10^6$	$7.11 \cdot 10^7$
	3	"	$9.27 \cdot 10^0$	$1.08 \cdot 10^9$	$7.11 \cdot 10^{10}$

$K(A,b)\|(\Delta A, \Delta b)\|_F / \|(A, b)\|_F \ll 1$, our forward error estimate based on the condition number (which is genuinely of first order nature) well approaches the true error, which shows the relevance of our approach. It is the topic of further studies to generalize this observation and explore the reliability of the first order analysis on a broader set of test problems.

5. Conclusion. We proposed sensitivity analysis tools for the TLS problem when the genericity condition is satisfied. We provided closed formulas for the condition number of a linear function of the TLS solution when the perturbations of data are measured normwise. We also described an algorithm based on an adjoint formula, and we expressed this condition number and an upper bound on it in terms of the SVDs of $[A, b]$ and A. We illustrated the use for these quantities in three numerical examples.

REFERENCES

[1] E. ANDERSON, Z. BAI, C. BISCHOF, S. BLACKFORD, J. DEMMEL, J. DONGARRA, J. DU CROZ, A. GREENBAUM, S. HAMMARLING, A. MCKENNEY, AND D. SORENSEN, *LAPACK Users' Guide*, 3rd ed., SIAM, Philadelphia, 1999.
[2] M. ARIOLI, M. BABOULIN, AND S. GRATTON, *A partial condition number for linear least-squares problems*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 413–433.

[3] M. Baboulin, J. Dongarra, S. Gratton, and J. Langou, *Computing the conditioning of the components of a linear least squares solution*, Numer. Linear Algebra Appl., 16 (2009), pp. 517–533.

[4] M. Baboulin and S. Gratton, *Using dual techniques to derive componentwise and mixed condition numbers for a linear function of a linear least squares solution*, BIT, 49 (2009), pp. 3–19.

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

[6] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, *ScaLAPACK Users' Guide*, SIAM, Philadelphia, 1997.

[7] S. Chandrasekaran and I. C. F. Ipsen, *On the sensitivity of solution components in linear systems of equations*, SIAM J. Matrix Anal. Appl., 16, 1995, pp. 93–112.

[8] A. J. Cox and N. J. Higham, *Accuracy and stability of the null space method for solving the equality constrained least squares problem*, BIT, 39 (1999), pp. 34–50.

[9] F. Cucker, H. Diao, and Y. Wei, *On mixed and componentwise condition numbers for Moore–Penrose inverse and linear least squares problems*, Math. Comp., 76 (2007), pp. 947–963.

[10] L. Eldén, *Perturbation theory for the least squares problem with linear equality constraints*, SIAM J. Numer. Anal., 17 (1980), pp. 338–350.

[11] A. J. Geurts, *A contribution to the theory of condition*, Numer. Math., 39 (1982), pp. 85–96.

[12] 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

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

[14] A. Graham, *Kronecker Products and Matrix Calculus with Applications*, Wiley, New York, 1981.

[15] S. Gratton, *On the condition number of linear least squares problems in a weighted Frobenius norm*, BIT, 36 (1996), pp. 523–530.

[16] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.

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

[18] I. C. F. Ipsen, *Numerical Matrix Analysis: Linear Systems and Least Squares*, SIAM, Philadelphia, 2009.

[19] E. D. Kaplan, *Understanding GPS: Principles and Applications*, Artech House Publishers, Boston, 1996.

[20] W. M. Kaula, *Theory of Satellite Geodesy*, Blaisdell Press, Waltham, MA, 1966.

[21] C. S. Kenney, A. J. Laub, and M. S. Reese, *Statistical condition estimation for linear least squares*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 906–923.

[22] C. Paige and Z. Strakoš, *Core problems in linear algebraic systems*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 861–875.

[23] J. Rice, *A theory of condition*, SIAM J. Numer. Anal., 3 (1966), pp. 3–287.

[24] G. W. Stewart, *Matrix Algorithms, Volume* II: *Eigensystems*, SIAM, Philadelphia, 2001.

[25] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1991.

[26] L. Zhou, L. Lin, Y. Wei, and S. Qiao, *Perturbation analysis and condition numbers of scaled total least squares problems*, Numer. Algorithms, 51 (2009), pp. 381–399.