UCLA

COMPUTATIONAL AND APPLIED MATHEMATICS

Perturbation Theory for Orthogonal Projection Methods with Applications to Least Squares and Total Least Squares

Ricardo D. Fierro
James R. Bunch

October 1992

(Revised March 1994)

CAM Report 92-44

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

PERTURBATION THEORY FOR ORTHOGONAL PROJECTION METHODS WITH APPLICATIONS TO LEAST SQUARES AND TOTAL LEAST SQUARES

RICARDO D. FIERRO* AND JAMES R. BUNCH†

Abstract. The stabilized versions of the least squares (LS) and total least squares (TLS) methods are two examples of orthogonal projection methods commonly used to "solve" the overdetermined system of linear equations $AX \approx B$ when A is nearly rank-deficient. In practice, when this system represents the noisy version of the rank-deficient, zero residual problem $A_0X_0 = B_0$, TLS usually yields a more accurate estimate of the exact solution. However, current perturbation theory does not justify the superiority of TLS over LS.

In this paper we establish a model for orthogonal projection methods by reformulating the parameter estimation problem as an equivalent problem of nullspace determination. When the method is based on the singular value decomposition of the matrix $[A\ B]$, the model specializes to the well known TLS method. We derive new lower and upper perturbation bounds for orthogonal projection methods in terms of the subspace angle, which shows how the perturbation of the approximate nullspace affects the accuracy of the solution. Using the new bounds, we show the TLS perturbation bound is generally smaller than the corresponding one for LS, which means TLS is usually more robust than LS under perturbations of all the data. Also, the bounds permit a comparison between the LS and TLS solutions, as well as for any two competing orthogonal projection methods. We include numerical simulations to substantiate our conclusions.

1. Introduction. In numerous applications one is faced with estimating the relationship between the columns of the data matrix $A \in \Re^{m\times n}$ and the observation matrix $B \in \Re^{m\times d}$ $(m \geq n+d)$ in the overdetermined system of linear equations

$$AX \approx B. \tag{1}$$

For example, in spectral estimation this problem must be solved in the noisy forward or forward-backward linear prediction technique for resolving closely spaced sinusoids [1],[20]. Usually (1) does not have a solution and compatibility must be restored by "fitting" $AX \approx B$ with a compatible system

$$\tilde{A}X = \tilde{B},$$

and the "correction" matrices are $\triangle\tilde{A} = A - \tilde{A}$ and $\triangle\tilde{B} = B - \tilde{B}$.

The least squares (LS) or total least squares (TLS) methods are commonly used for (1). These methods, as well as the analysis pertaining to them, are often based on the singular value decomposition (SVD). Denote the SVD of A (cf. [7, p.70]) in the dyadic form by

$$A = \sum_{i=1}^{n} \sigma_i' u_i' v_i'^T, \tag{2}$$

where $\sigma_1' \geq \sigma_2' \geq \cdots \geq \sigma_n' \geq 0$ and the u_i''s, as well as the v_i''s, are mutually orthogonal.

Literature on LS is abundant (e.g., [2],[5],[6],[9]). If the numerical rank of A is $k < n$, it is well known that small perturbations in A or B may cause disproportionately large changes in the ordinary LS solution $X_{OLS} = \sum_{i=1}^{n} v_i' \frac{(u_i'^T B)}{\sigma_i'}$. The *truncated* LS method stabilizes the solution by solving the related LS problem

$$A_k X \approx B, \tag{3}$$

where $A_k = \sum_{i=1}^{k} \sigma_i' u_i' v_i'^T$. Note that the approximation of A by A_k is independent of B and $\|\triangle A_k\| = \sigma_{k+1}'$, where $\|\cdot\| \equiv \|\cdot\|_2$ denotes the Euclidean norm, unless otherwise stated. Here,

* Department of Mathematics, California State University, San Marcos, CA 92096 (fierro@thunder.csusm.edu).
† Department of Mathematics, UCSD, La Jolla, CA 92093-0112 (jbunch@ucsd.edu).

one attempts to achieve stability in the solution at the expense of a slightly larger residual. A perturbation analysis for the truncated LS method for the case $d = 1$ is given by Hansen in [8]. Let

$$X_{LS} = \sum_{i=1}^{k} v_i' \frac{(u_i'^T B)}{\sigma_i'} \tag{4}$$

denote the minimum norm LS solution to (3). The set S_{LS} of all solutions is characterized by $S_{LS} = \{X | X = A_k^\dagger B + (I - A_k^\dagger A_k)Z, \quad \forall Z \in \Re^{n \times d}\}$.

Equation (1) often contains independently and identically distributed errors in both A and B. Denote the SVD of $[A\ B]$ in the dyadic form by

$$[A\ B] = \sum_{i=1}^{n+d} \sigma_i u_i v_i^T \tag{5}$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{n+d}$ and the u_i's, as well as the v_i's, are mutually orthogonal. In the setting of orthogonal projection methods, Van der Sluis and Veltkamp [13] provided brief results about the TLS approach to (1). Golub and Van Loan [6] provided the first analysis of TLS ($k = n, d = 1$) using the SVD, and later the results were extended ($k = n, d \geq 1$) by Van Huffel and Vandewalle, who also examined the case where A is exactly rank k [14],[15].

In TLS we consider a perturbation of A which depends on B, as follows:

$$\text{minimize} \quad \|[A\ B] - [C\ D]\| \text{ subject to } \quad CX = D.$$

For stability reasons (more explicitly, see TLS perturbation theory in §3) when the numerical rank of A is $k \leq n$, the TLS problem is reformulated as finding the minimizer of the constrained optimization problem

$$\text{minimize} \qquad \|[A\ B] - [C\ D]\| \tag{6}$$

$$\text{subject to} \quad \text{rank } (C) = \text{ rank}([C\ D]) = k, \tag{7}$$

provided such a minimizer exists; this will be assumed in this section, and enforced in the following sections with a mild condition.

A nearest rank-k matrix approximation to $[A\ B]$, given by

$$[\hat{A}\ \hat{B}] = \sum_{i=1}^{k} \sigma_i u_i v_i^T, \tag{8}$$

is the most likely candidate for the solution. $\|\triangle \hat{A}\|$ is the error in the approximation of A by $\hat{A}$ and $\|[\triangle \hat{A}\ \triangle \hat{B}]\| = \sigma_{k+1}$. The *truncated* TLS solution, X_{TLS}, is the minimum norm solution to

$$\hat{A}X = \hat{B}\,. \tag{9}$$

As proven in [13], the set S_{TLS} of solutions to (9) is the same as the set of LS solutions to $\hat{A}X \approx B$. Thus, $S_{TLS} = \{X | X = \hat{A}^\dagger B + (I - \hat{A}^\dagger \hat{A})Z, \quad \forall Z \in \Re^{n \times d}\}$.

The purpose of this paper is to develop theory which explains the superiority of (truncated) TLS over (truncated) LS observed empirically in the literature. We shall assume that A is possibly near-rank deficient (in which case the truncated versions of these methods are employed), and the results and conclusions are equally valid if $\text{rank}(A) = n$.

The paper is organized as follows. In §2 we establish a model for a general orthogonal projection method M, such as LS and TLS, by reformulating the parameter estimation problem as an equivalent problem of nullspace determination. When the method is based on the SVD of the matrix

$[A\ B]$, the model specializes to the well known TLS method. For an arbitrary orthogonal projection method M, we derive lower and upper perturbation bounds for the solutions. The bounds are in terms of the subspace angle between approximate nullspaces, which shows how the perturbation of the approximate nullspace affects the accuracy of the solution. The perturbation result is briefly mentioned in [4] but is formulated and explored more thoroughly in this paper. In §3 we apply the general perturbation bounds to LS and TLS. In situations where TLS typically applies, our numerical experience indicates the TLS bounds are usually smaller than the corresponding LS bounds. This means TLS is usually superior to LS under perturbations of all the data. In §4 we examine the difference between the LS and TLS solution using the general perturbation bounds as a platform. The bounds actually allow for a comparison between the solutions of any two competing orthogonal projection methods, such as LS or TLS and a rank revealing QR factorization [3],[4]. Numerical simulations are included in §5 to illustrate the conclusions. Finally, we summarize our conclusions in §6.

At this point we introduce notations and definitions used in this paper. Superscripts T and $\dagger$ denote the transpose and Moore-Penrose pseudoinverse of a matrix, respectively. Let $\mathcal{R}(D)$ and $\mathcal{N}(D)$ denote the range and kernel of the matrix D. Lower case Greek letters are scalars.

Definition [6, p.75] Given $D \in \Re^{p\times q}$, $P \in \Re^{p\times p}$ is the *orthogonal projection* onto $\mathcal{R}(D)$ if $\mathcal{R}(P) = \mathcal{R}(D)$, $P^2 = P$, and $P^T = P$.

For LS, $[A_k\ B_k] = P[A\ B]$ where $P = [u_1', ..., u_k'][u_1', ..., u_k']^T$ and $\mathcal{R}(P) = \mathcal{R}(A_k)$. For TLS, $[\hat{A}\ \hat{B}] = P[A\ B]$ where $P = [u_1, ..., u_k][u_1, ..., u_k]^T$ and $\mathcal{R}(P) = \mathcal{R}(\hat{A})$. Now, let $P_D \equiv DD^\dagger$.

Definition [18] The matrix C is an *acute* perturbation of D if $\|P_C - P_D\| < 1$ and $\|P_{C^T} - P_{D^T}\| < 1$. We also say that C and D are acute.

Finally, the next definition allows one to compare two subspaces.

Definition [6, p.76] Suppose $\mathcal{R}(C)$ and $\mathcal{R}(D)$ are equidimensional subspaces of $\Re^{q\times q}$. We define the ***distance*** between these two subspaces by $\sin\phi \equiv \|P_C - P_D\|$, where ϕ is the largest angle between the two subspaces.

2. General Model and Perturbation Bounds.

In this section we establish a model for a general orthogonal projection method M, such as LS or TLS, by reformulating the parameter estimation problem as an equivalent problem of nullspace determination and by providing lower and upper perturbation bounds for the solutions. The bounds apply to any parameter estimation method which can be reformulated as an equivalent problem of nullspace determination. The key is to determine an orthonormal basis Y for the nullspace of a compatible system which approximates $AX \approx B$ (or a basis for an approximate nullspace of $[A\ B]$) followed by a change of basis. This information can be readily extracted, for example, from the SVD, URV, or ULV decomposition, and indirectly from a rank revealing QR decomposition of $[A\ B]$.

As shown by [3],[4] a rank revealing QR factorization can be used to compute a solution based on an approximate nullspace. Stewart [11] proposed a method which employs a two-sided orthogonal decomposition to determine an approximate nullspace of a matrix (for subspace tracking). An efficient implementation of the ULV or URV algorithm to compute a TLS solution is proposed in [17]. The choice is usually governed by the desired accuracy and the computational and implementation requirements.

We shall now present the model. Denote by X_M the minimum norm solution to

$$\tilde{A}X = \tilde{B} \tag{10}$$

where $[\tilde{A}\ \tilde{B}]$ is a rank-k matrix approximation to $[A\ B]$ based on method M, and $[\triangle\tilde{A}\ \triangle\tilde{B}] = [A\ B] - [\tilde{A}\ \tilde{B}]$. The following result gives a sufficient condition for the existence of X_M, i.e., the compatibility of (10).

THEOREM 2.1. *Let A have the SVD as in (2), $P \in \Re^{m\times m}$ an orthogonal projection matrix, $[\tilde{A}\ \tilde{B}] = P[A\ B]$, and $\triangle\tilde{A} = A - \tilde{A}$. Then $\tilde{A}X = \tilde{B}$ is compatible provided $||\triangle\tilde{A}|| < \sigma'_k$.*
Proof: See [4].

This condition means that $\tilde{A}$ must be an acute perturbation of A_k. The solution set S_M to (10) is characterized by

$$S_M = \{X | X = \tilde{A}^\dagger B + (I - \tilde{A}^\dagger \tilde{A})Z, \quad \forall Z \in \Re^{n\times d}\}. \tag{11}$$

Let $[\bar{A}\ \bar{B}] = [A\ B] + [\triangle A\ \triangle B]$ represent a perturbation of $[A\ B]$. Let $[\breve{A}\ \breve{B}]$ denote a rank-k matrix approximation to $[\bar{A}\ \bar{B}]$ based on method M and define $[\triangle\breve{A}\ \triangle\breve{B}] = [\bar{A}\ \bar{B}] - [\breve{A}\ \breve{B}]$. We shall always assume the mild condition $||\triangle\breve{A}|| + ||\triangle A|| < \sigma'_k$. By Theorem 2.1, $\breve{A}X = \breve{B}$ is compatible. Denote by $\bar{X}_M$ the minimum norm solution to

$$\breve{A}X = \breve{B}. \tag{12}$$

Now we are prepared to develop our model. From (10) we have

$$[\tilde{A}\ \tilde{B}] \begin{bmatrix} X_M \\ -I_d \end{bmatrix} = 0.$$

Now, let the columns of $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$, with $Y_1 \in \Re^{n\times(n-k+d)}$ and $Y_2 \in \Re^{d\times(n-k+d)}$, form an orthonormal basis for the kernel of $[\tilde{A}\ \tilde{B}]$, denoted $\mathcal{N}([\tilde{A}\ \tilde{B}])$. Since

$$\mathcal{R}\left(\begin{bmatrix} X_M \\ -I_d \end{bmatrix}\right) = \mathcal{R}\left(\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}\right),$$

it follows

$$\begin{bmatrix} X_M \\ -I_d \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} S_Y, \tag{13}$$

for some $S_Y \in \Re^{(n-k+d)\times d}$. This yields the underdetermined compatible system

$$Y_2 S_Y = -I_d. \tag{14}$$

The general solution set is $\{S_Y | S_Y = -Y_2^\dagger + Z, \quad \forall Z \in \mathcal{N}(Y_2)\}$. Taking the norm of both sides of (13) we get

$$\sqrt{1 + ||X_M||^2} = ||S_Y||.$$

Since X_M is the minimum norm solution, we take $Z = 0$ to get $S_Y = -Y_2^\dagger$. Therefore,

$$\begin{aligned} X_M &= -Y_1 Y_2^\dagger \\ &= -Y_1 Y_2^T (Y_2 Y_2^T)^{-1}. \end{aligned}$$

It can also be shown that if the columns of $Z = [Z_1^T\ Z_2^T]^T$ are orthonormal and span the orthogonal complement of $\mathcal{R}(Y)$, then

$$X_M = Z_1 Z_2^T (I_d - Z_2 Z_2^T)^{-1}.$$

In addition, if $Q \in \Re^{(n-k+d)\times(n-k+d)}$ is an orthogonal matrix such that

$$\begin{bmatrix} Y_1Q \\ Y_2Q \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} Q = YQ = \begin{array}{c} \begin{matrix} n-k & d \end{matrix} \\ \begin{bmatrix} D & C \\ 0 & \Gamma \end{bmatrix} \end{array} \begin{matrix} n \\ d, \end{matrix}$$

then it follows

$$\begin{aligned} X_M = -Y_1Y_2^\dagger &= -(Y_1Q)(Y_2Q)^\dagger \\ &= -[D\,C]\,[0\,\Gamma]^\dagger \\ &= -[D\,C]\begin{bmatrix} 0 \\ \Gamma^{-1} \end{bmatrix} \\ &= -C\,\Gamma^{-1}, \end{aligned} \tag{15}$$

and one only needs to solve a triangular system of equations to compute X_M.

If M $\equiv$ TLS, then this model specializes to the well known computation of the TLS solution by Golub and Van Loan [6] for $k = n$, Van Huffel and Vandewalle [15], and Zoltowski [21]. We remark that for $k = n$ the Generalized CS Theorem [10] can be invoked to develop relationships similar to those presented in [16, Ch. 3] for the TLS model.

Now, if $Z = [Z_1^T\ Z_2^T]^T$ denotes another orthonormal basis for $\mathcal{N}([\tilde{A}\ \tilde{B}])$ then the "orthogonal Procrustes" problem for Y and Z can be solved with zero residual, i.e., there exists an orthogonal matrix Q satisfying $Y = ZQ$. Then

$$X_M = -Y_1\,Y_2^\dagger = -Y_1QQ^TY_2^\dagger = (-Y_1Q)\,(Y_2Q)^\dagger = -Z_1\,Z_2^\dagger,$$

which means choice of basis is unimportant and the solution is completely determined by the kernel of $[\tilde{A}\ \tilde{B}]$.

Assume the mild condition $\max(\,\|\triangle\tilde{A}\|, \|\triangle\breve{A}\| + \|\triangle A\|\,) < \sigma_k'$ is satisfied and let the columns of $\bar{Y}$ denote an orthonormal basis for $\mathcal{N}([\breve{A}\ \breve{B}])$. Let $\bar{Q} \in \Re^{(n-k+d)\times(n-k+d)}$ be an orthogonal matrix such that

$$\begin{bmatrix} \bar{Y}_1\bar{Q} \\ \bar{Y}_2\bar{Q} \end{bmatrix} = \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \end{bmatrix} \bar{Q} = \bar{Y}\bar{Q} = \begin{array}{c} \begin{matrix} n-k & d \end{matrix} \\ \begin{bmatrix} \bar{D} & \bar{C} \\ 0 & \bar{\Gamma} \end{bmatrix} \end{array} \begin{matrix} n \\ d. \end{matrix}$$

Then

$$\begin{aligned} \bar{X}_M &= -\bar{Y}_1\,\bar{Y}_2^\dagger \\ &= -\bar{C}\,\bar{\Gamma}^{-1}. \end{aligned} \tag{16}$$

Thus,

$$\begin{bmatrix} \bar{X}_M \\ -I_d \end{bmatrix} - \begin{bmatrix} X_M \\ -I_d \end{bmatrix} = \begin{bmatrix} C \\ \Gamma \end{bmatrix}\Gamma^{-1} - \begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix}\bar{\Gamma}^{-1}. \tag{17}$$

Let $W \in \Re^{(n+d)\times n}$ denote an orthonormal matrix with the partition

$$W = \begin{array}{c} n \\ \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \end{array} \begin{matrix} n \\ d \end{matrix}$$

such that $W^TY = 0$ (i.e., W "completes the space"). From equation (17) it follows

$$W^T\left(\begin{bmatrix} \bar{X}_M \\ -I_d \end{bmatrix} - \begin{bmatrix} X_M \\ -I_d \end{bmatrix}\right) = -W^T\begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix}\bar{\Gamma}^{-1}, \tag{18}$$

and consequently

$$(19) \qquad W_1^T(\bar{X}_M - X_M) = -W^T \begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix} \bar{\Gamma}^{-1}.$$

Note that

$$\sin \phi_M \equiv \left\| W^T \begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix} \right\|$$

denotes the sine of the largest subspace angle between $\mathcal{R}\left(\begin{bmatrix} C \\ \Gamma \end{bmatrix}\right)$ and $\mathcal{R}\left(\begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix}\right)$. We are now ready to present the main result of this section.

Theorem 2.2. *Denote $[\tilde{A}\ \tilde{B}] = [A\ B] + [\Delta A\ \Delta B]$. Let X_M and $\bar{X}_M$ denote the minimum norm solutions to the compatible systems $\tilde{A}X = \tilde{B}$ and $\check{A}X = \check{B}$ obtained from method M as in (10) and (12), provided $max(\|\Delta\tilde{A}\|, \|\Delta\check{A}\| + \|\Delta A\|) < \sigma_k'$. Then*

$$(20) \qquad \sin \phi_M \ \leq \ \|X_M - \bar{X}_M\| \ \leq \ \sin \phi_M \sqrt{1 + \|X_M\|^2}\sqrt{1 + \|\bar{X}_M\|^2},$$

where ϕ_M denotes the largest subspace angle between $\mathcal{R}\left(\begin{bmatrix} C \\ \Gamma \end{bmatrix}\right)$ and $\mathcal{R}\left(\begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix}\right)$.

Proof: From the CS Theorem [10], since

$$\begin{bmatrix} W_1 & C \\ W_2 & \Gamma \end{bmatrix}$$

is an orthogonal matrix where W_1 and Γ are square matrices, we know $\sigma_{\min}^{-1}(W_1) = \|\Gamma^{-1}\|$.

From (19) we have

$$\begin{aligned} \sigma_{\min}(W_1)\,\|\bar{X}_M - X_M\| &\leq \|W_1^T(\bar{X}_M - X_M)\| \\ &= \left\| W^T \begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix} \bar{\Gamma}^{-1} \right\| \\ &\leq \sin \phi_M \|\bar{\Gamma}^{-1}\|, \end{aligned}$$

or

$$\|\bar{X}_M - X_M\| \ \leq \ \sin \phi_M \|\bar{\Gamma}^{-1}\|\,\sigma_{\min}^{-1}(W_1).$$

Hence, it follows

$$\begin{aligned} \|\bar{X}_M - X_M\| &\leq \sin \phi_M \|\bar{\Gamma}^{-1}\|\,\|\Gamma^{-1}\| \\ &= \sin \phi_M \sqrt{1 + \|\bar{X}_M\|^2}\sqrt{1 + \|X_M\|^2} \end{aligned}$$

and this proves the upper bound. To prove the lower bound,

$$\begin{aligned} \|\bar{X}_M - X_M\| &\geq \|W_1^T(\bar{X}_M - X_M)\| \\ &= \left\| W^T \begin{bmatrix} \bar{C} \\ \bar{\Gamma} \end{bmatrix} \bar{\Gamma}^{-1} \right\| \\ &\geq \sin \phi_M\ \sigma_{\min}(\bar{\Gamma}^{-1}) \\ &\geq \sin \phi_M, \end{aligned}$$

since $\sigma_{\min}(\bar{\Gamma}^{-1}) \geq 1$. This completes the proof.

We remark that Theorem 2.2 may also be applied to the situation where X_M and $\bar{X}_M$ are determined by different orthogonal projection methods; we pursue this matter in §4 only for the case $\|\Delta A\| = 0$.

Theorem 2.2 shows that for method M every perturbation $[\Delta A\ \Delta B]$ which makes $\mathcal{N}([\breve{A}\ \breve{B}])$ equal to $\mathcal{N}([\tilde{A}\ \tilde{B}])$ makes $\breve{A}X = \breve{B}$ compatible and yields the same solution X_M, since $\sin\phi_M = 0$. The perturbations $[\Delta A\ \Delta B]$ which achieve this are characterized by

$$[\Delta A\ \Delta B] = -[\Delta\tilde{A}\ \Delta\tilde{B}] + H^T \tag{21}$$

where $\mathcal{R}(H) \perp \mathcal{R}(Y)$. Thus various perturbations may produce the same solution. For arbitrary $[\Delta A\ \Delta B]$ we conclude that for method M the perturbation effect depends upon the noise distribution in $\mathcal{N}([\breve{A}\ \breve{B}])$. Van Huffel and Vandewalle [16, p.190] reached the same conclusion (21) for the perturbation effects in the TLS problem.

Theorem 2.2 also shows as soon as the kernel is perturbed such that $0 < \sin\phi_M$, the new solution $\bar{X}_M$ cannot coincide with X_M. In the presence of noise $[\Delta A\ \Delta B]$, the best accuracy an orthogonal projection method M can achieve is measured by $\sin\phi_M$. Finally, the $\sqrt{\cdot}$ in the perturbation bound is actually a "condition number" for the orthogonal projection method, cf. [4].

Hereto we denote

$$\sin\theta_M = \left\| YY^T - \bar{Y}\bar{Y}^T \right\|,$$

the sine of the largest angle between $\mathcal{N}([\tilde{A}\ \tilde{B}])$ and $\mathcal{N}([\breve{A}\ \breve{B}])$. Although the columns of $[C^T\ \Gamma^T]^T$ and $[\bar{C}^T\ \bar{\Gamma}^T]^T$ are nullvectors of $[\tilde{A}\ \tilde{B}]$ and $[\breve{A}\ \breve{B}]$, respectively, generally $\sin\phi_M \neq \sin\theta_M$.

A straightforward approach for finding an upper bound for $\sin\phi_M$ in terms of familiar parameters appears difficult, hence we approach the problem by first finding an upper bound for $\sin\theta_M$. There are several reasons why we might be interested in an upper bound for $\sin\theta_M$.

- First, $\sin\phi_M = \sin\theta_M$ whenever $k = n$, a situation that arises in many applications.
- Second, an upper bound for $\sin\theta_M$ in terms of familiar parameters is feasible.
- Third, we can bound $\sin\phi_M$ by $\sin\theta_M$ plus a term, as follows.

Define the projection matrices $P_Y = YY^T$, $P_{\bar{Y}} = \bar{Y}\bar{Y}^T$, $P_{C,\Gamma} = [C^T\ \Gamma^T]^T[C^T\ \Gamma^T]$, $P_{\bar{C},\bar{\Gamma}} = [\bar{C}^T\ \bar{\Gamma}^T]^T[\bar{C}^T\ \bar{\Gamma}^T]$, $P_D = [D^T\ 0]^T[D^T\ 0]$, and $P_{\bar{D}} = [\bar{D}^T\ 0]^T[\bar{D}^T\ 0]$. Using the facts $P_Y = P_D + P_{C,\Gamma}$, $P_{\bar{Y}} = P_{\bar{D}} + P_{\bar{C},\bar{\Gamma}}$, $P_D P_{C,\Gamma} = 0$, and $P_{\bar{D}} P_{\bar{C},\bar{\Gamma}} = 0$, then for any vector $z \in \Re^{n+d}$

$$\begin{aligned}\|(P_Y - P_{\bar{Y}})z\|^2 &= \|(P_D - P_{\bar{D}})z\|^2 \\ &+ \|(P_{C,\Gamma} - P_{\bar{C},\bar{\Gamma}})z\|^2 - z^T(P_D P_{\bar{C},\bar{\Gamma}} + P_{\bar{D}} P_{C,\Gamma} + P_{C,\Gamma} P_{\bar{D}} + P_{\bar{C},\bar{\Gamma}} P_D)z.\end{aligned}$$

In particular, if z_Γ is a unit vector such that $\|P_{C,\Gamma} - P_{\bar{C},\bar{\Gamma}}\| = \|(P_{C,\Gamma} - P_{\bar{C},\bar{\Gamma}})z_\Gamma\|$, then

$$\sin\phi_M \leq \sin\theta_M + \epsilon_M,$$

where

$$\epsilon_M = \left| {z_\Gamma}^T(P_D P_{\bar{C},\bar{\Gamma}} + P_{\bar{D}} P_{C,\Gamma} + P_{C,\Gamma} P_{\bar{D}} + P_{\bar{C},\bar{\Gamma}} P_D)z_\Gamma - \|(P_D - P_{\bar{D}})z_\Gamma\|^2 \right|^{\frac{1}{2}}.$$

Note $\epsilon_M = 0$ whenever $k = n$.

Our numerical experiments in §5 suggest $\epsilon_M \leq \sin\theta_M$, and experimentally we observed $\sin\phi_M \leq 2\sin\theta_M$. Although there may exist counterexamples, we did not encounter any.

The abovementioned reasons motivate us to bound $\sin\theta_M$ in terms of the fittings $[\tilde{A}\ \tilde{B}]$ and $[\breve{A}\ \breve{B}]$, corrections, and the perturbation. First, we need to state some useful results.

LEMMA 2.3. *[13] Let $[C\ D] = [\tilde{C}\ \tilde{D}] + [\Delta\tilde{C}\ \Delta\tilde{D}]$ where $[\tilde{C}\ \tilde{D}]$ is obtained by orthogonal projection. Then $[\tilde{C}\ \tilde{D}]^\dagger[\Delta\tilde{C}\ \Delta\tilde{D}] = 0$.*

LEMMA 2.4. *[18, Theorem 3.14] If C is an acute perturbation of D, with $C = D + E$, then $\|C^\dagger - D^\dagger\| \leq \mu\|C^\dagger\|\,\|D^\dagger\|\,\|E\|$, where $\mu = (1+\sqrt{5})/2$.*

We are ready to derive an upper bound for $\sin\theta_M$.

THEOREM 2.5. *Let $[\bar{A}\ \bar{B}] = [A\ B] + [\Delta A\ \Delta B]$ with $[\tilde{A}\ \tilde{B}]$ and $[\check{A}\ \check{B}]$ obtained as in (10) and (12) using method M. Define $\mu = (1+\sqrt{5})/2$. If $\max(\|\Delta\tilde{A}\|, \|\Delta\check{A}\| + \|\Delta A\|) < \sigma_k'$ then*

$$\sin\theta_M \le \|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\Delta A\ \Delta B]\| + \mu\,\|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\check{A}\ \check{B}]^\dagger\|\,\|[\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}]\|\,\|\Delta\check{A}\ \Delta\check{B}\|$$

where θ_M is the largest subspace angle between $\mathcal{N}([\tilde{A}\ \tilde{B}])$ and $\mathcal{N}([\check{A}\ \check{B}])$.

Proof: Let $\check{R}^\perp \equiv I - [\check{A}\ \check{B}]^\dagger[\check{A}\ \check{B}]$. Then

$$\begin{aligned}
\sin\theta_M &= \|[\tilde{A}\ \tilde{B}]^\dagger[\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}]^\dagger[\check{A}\ \check{B}]\| \\
&= \|[\tilde{A}\ \tilde{B}]^\dagger[\tilde{A}\ \tilde{B}]\check{R}^\perp\| \\
&= \|[\tilde{A}\ \tilde{B}]^\dagger([\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}])\check{R}^\perp\| \\
&\le \|[\tilde{A}\ \tilde{B}]^\dagger([\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}])\| \\
&= \|[\tilde{A}\ \tilde{B}]^\dagger([\Delta A\ \Delta B] - [\Delta\check{A}\ \Delta\check{B}])\| \\
&= \|[\tilde{A}\ \tilde{B}]^\dagger[\Delta A\ \Delta B] - ([\tilde{A}\ \tilde{B}]^\dagger - [\check{A}\ \check{B}]^\dagger)[\Delta\check{A}\ \Delta\check{B}]\| \\
&\le \|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\Delta A\ \Delta B]\| + \|[\tilde{A}\ \tilde{B}]^\dagger - [\check{A}\ \check{B}]^\dagger\|\,\|[\Delta\check{A}\ \Delta\check{B}]\| \\
&\le \|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\Delta A\ \Delta B]\| + \mu\,\|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\check{A}\ \check{B}]^\dagger\|\,\|[\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}]\|\,\|[\Delta\check{A}\ \Delta\check{B}]\|.
\end{aligned}$$

The last inequality follows from Lemma 2.4. This completes the proof.

This proof follows a similar line of reasoning as an argument by van der Sluis and Veltkamp [13]. Note that in Theorem 2.5

$$\mu\,\|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\check{A}\ \check{B}]^\dagger\|\,\|[\tilde{A}\ \tilde{B}] - [\check{A}\ \check{B}]\|\,\|[\Delta\check{A}\ \Delta\check{B}]\| \;<<\; \|[\tilde{A}\ \tilde{B}]^\dagger\|\,\|[\Delta A\ \Delta B]\|$$

provided $[\Delta A\ \Delta B]$ is not too large and (1) is not too incompatible. This is illustrated in the next section when we see how Theorems 2.2 and 2.5 may be used to derive perturbation theory for the popular orthogonal projection methods LS and TLS.

3. LS and TLS Perturbation Bounds.

In this section we use the general perturbation bounds in §2 to derive new upper perturbation bounds for LS and TLS by bounding the subspace angle. In §3.1 we examine the special case $\text{rank}(A) = k$ and (1) is compatible. We show that the TLS subspace angle is usually lower than the corresponding LS subspace angle (assuming ϵ_{TLS} and ϵ_{LS} are sufficiently small), which explains the superiority of TLS over LS observed by many in various applications (e.g., see [16] and the references cited therein). In §3.2 we will investigate the more general case when A is nearly rank-deficient and (1) is incompatible.

3.1. Bounds when Rank(A) = k and $AX = B$.

In many sinusoidal frequency estimations problems (e.g., see [12],[20]) the coefficient matrix A has rank $k \le n$ and $Ax = b$ ($d = 1$) is compatible in the absence of white noise. The elements of the solution x are the coefficients of a polynomial

$$P(z^{-1}) = 1 - \sum_{i=1}^{n} x_i z^{-i}$$

and the true frequencies are determined from the angular position of the roots of $P(z)$ on the unit circle. However, due to the presence of noise, one has to estimate the polynomial coefficients from a perturbed problem $\bar{A}x \approx \bar{b}$ and consequently this corrupts the computed frequencies.

For now we shall assume $\text{rank}(A) = k$ and (1) is compatible, hence the LS and TLS solutions to (1) coincide: $X_0 = X_{LS} = X_{TLS}$. Letting $[\bar{A}\ \bar{B}] = [A\ B] + [\Delta A\ \Delta B]$, denote by $\bar{A}_k$ the nearest rank-k approximation to $\bar{A} = \bar{A}_k + \Delta\bar{A}_k$. We must solve the truncated LS problem $\bar{A}_k X \approx \bar{B}$. This is equivalent to finding the minimum norm solution $\bar{X}_{LS}$ to the compatible system

$$\bar{A}_k X = \bar{B}_k, \tag{22}$$

where $\bar{B}_k = \bar{A}_k \bar{A}_k^\dagger \bar{B}$ is the orthogonal projection of $\bar{B}$. From the results in §2, it remains to bound the subspace angle $\sin\phi_{LS}$, which in turn requires an upper bound for $\sin\theta_{LS}$:

$$\sin\theta_{LS} = \operatorname{dist}(\mathcal{N}([A\ B])\,, \mathcal{N}([\bar{A}_k\ \bar{B}_k])\,).$$

We are also interested in solving $\bar{A}X \approx \bar{B}$ in the truncated TLS sense. Let $[\breve{A}\ \breve{B}] = \sum_{i=1}^{k} \bar{\sigma}_i \bar{u}_i \bar{v}_i^T$ denote the nearest rank-k approximation to $[\bar{A}\ \bar{B}] = \sum_{i=1}^{n+d} \bar{\sigma}_i \bar{u}_i \bar{v}_i^T$. Then TLS finds the minimum norm solution $\bar{X}_{TLS}$ to the compatible system

$$\breve{A}X = \breve{B}. \tag{23}$$

In a similar manner as above, it remains to bound the subspace angle $\sin\phi_{TLS}$, which in turn requires an upper bound $\sin\theta_{TLS}$:

$$\sin\theta_{TLS} = \operatorname{dist}(\mathcal{N}([A\ B]), \mathcal{N}([\breve{A}\ \breve{B}])\,).$$

THEOREM 3.1. *Assume the rank of A is k and let X_0 denote the solution to the compatible system $AX = B$. Let $[\bar{A}\ \bar{B}] = [A\ B] + [\Delta A\ \Delta B]$, and denote by $\bar{X}_{LS}$ and $\bar{X}_{TLS}$ the minimum norm solutions to (22) and (23), respectively, provided $\|[\Delta A\ \Delta B]\| < \sigma_k'$. Then the following perturbation bounds hold:*

$$\begin{aligned} \sin\phi_{LS} \le \quad \|X_0 - \bar{X}_{LS}\| &\le \sin\phi_{LS}\sqrt{1+\|X_0\|^2}\,\sqrt{1+\|\bar{X}_{LS}\|^2} \\ &\le \left(\frac{\|[\Delta A\ \Delta B]\|}{\sigma_k' - \|\Delta A\|} + \epsilon_{LS}\right)\sqrt{1+\|X_0\|^2}\,\sqrt{1+\|\bar{X}_{LS}\|^2} \end{aligned}$$

$$\begin{aligned} \sin\phi_{TLS} \le \quad \|X_0 - \bar{X}_{TLS}\| &\le \sin\phi_{TLS}\sqrt{1+\|X_0\|^2}\,\sqrt{1+\|\bar{X}_{TLS}\|^2} \\ &\le \left(\frac{\|[\Delta A\ \Delta B]\|}{\sigma_k - \|[\Delta A\ \Delta B]\|} + \epsilon_{TLS}\right)\sqrt{1+\|X_0\|^2}\,\sqrt{1+\|\bar{X}_{TLS}\|^2}, \end{aligned}$$

Proof: From Theorem 2.2 and §2, $\sin\phi_{LS} \le \sin\theta_{LS} + \epsilon_{LS}$ and $\sin\phi_{TLS} \le \sin\theta_{TLS} + \epsilon_{TLS}$, so it remains to bound $\sin\theta_{TLS}$ and $\sin\theta_{LS}$. The assumptions imply $\|[\Delta\hat{A}\ \Delta\hat{B}]\| = \|[\Delta A_k\ R_k]\| = 0$, where $R_k = B - AX_k$. We could determine a bound using Theorem 3.2; however, under the conditions of this theorem, it is possible to eliminate a term bounding the subspace angle. In direct analogy to Theorem 2.5,

$$\begin{aligned} \sin\theta_{LS} &= \|[A\ B]^\dagger[A\ B] - [\bar{A}_k\ \bar{B}_k]^\dagger[\bar{A}_k\ \bar{B}_k]\| \\ & \|[A_k\ B_k]^\dagger[A_k\ B_k] - [\bar{A}_k\ \bar{B}_k]^\dagger[\bar{A}_k\ \bar{B}_k]\| \\ &\le \|[\bar{A}_k\ \bar{B}_k]^\dagger([A_k\ B_k] - [\bar{A}_k\ \bar{B}_k])\| \\ &= \|[\bar{A}_k\ \bar{B}_k]^\dagger\,([\Delta A\ \Delta B] + [\Delta A_k\ R_k])\| \\ &= \|[\bar{A}_k\ \bar{B}_k]^\dagger[\Delta A\ \Delta B]\| \\ &\le \|[\bar{A}_k\ \bar{B}_k]^\dagger\|\,\|[\Delta A\ \Delta B]\| \\ &\le \frac{\|[\Delta A\ \Delta B]\|}{\sigma_k' - \|\Delta A\|}. \end{aligned}$$

A similar argument will show that

$$\begin{aligned}\sin\theta_{TLS} &= \|[\breve{A}\ \breve{B}]^\dagger[\breve{A}\ \breve{B}] - [A\ B]^\dagger[A\ B]\| \\ &= \|[\breve{A}\ \breve{B}]^\dagger[\breve{A}\ \breve{B}] - [\hat{A}\ \hat{B}]^\dagger[\hat{A}\ \hat{B}]\| \\ &\le \|[\breve{A}\ \breve{B}]^\dagger([\hat{A}\ \hat{B}] - [\breve{A}\ \breve{B}])\| \\ &= \|[\breve{A}\ \breve{B}]^\dagger([\Delta A\ \Delta B] + [\Delta\hat{A}\ \Delta\hat{B}])\| \\ &= \|[\breve{A}\ \breve{B}]^\dagger[\Delta A\ \Delta B]\| \\ &\le \|[\breve{A}\ \breve{B}]^\dagger\|\,\|[\Delta A\ \Delta B]\| \\ &\le \frac{\|[\Delta A\ \Delta B]\|}{\sigma_k - \|[\Delta A\ \Delta B]\|}.\end{aligned}$$

Thus the desired results follow from Theorem 2.2 and §2. This completes the proof.

In our numerical simulations in §5, $\sin\phi_{TLS}$ and $\sin\theta_{TLS}$ are usually less than $\sin\phi_{LS}$ and $\sin\theta_{LS}$, respectively, hence TLS usually produces a more accurate estimate of the true solution, X_0, to the zero-residual, rank-deficient problem $AX = B$. Due to the tightness of the bounds, TLS usually produces a more accurate estimate than LS because the TLS subspaces are less sensitive to noise.

Both techniques are enhanced by increasing the respective k^{th} singular value, which in turn decreases the sensitivity of its subspace to perturbations. The results in Tables 1 and 3 illustrate this phenomenon. When $\frac{\|[\Delta A\ \Delta B]\|}{\sigma'_k} << 1$ then both techniques produce similar solutions.

3.2. Bounds when Numerical Rank(A) = k and $AX \approx B$. Now we wish to find perturbation bounds under the more general conditions (1) is incompatible and the numerical rank of A is k. Again, we turn to finding an upper bound for $\sin\theta_{TLS}$ and $\sin\theta_{LS}$.

Letting $[\bar{A}\ \bar{B}] = [A\ B] + [\Delta A\ \Delta B]$, denote by $\bar{A}_k$ the nearest rank-k approximation to $\bar{A} = \bar{A}_k + \Delta\bar{A}_k$. Then we must solve the problems

$$A_kX \approx B \quad \text{and} \quad \bar{A}_kX \approx \bar{B}$$

in the LS sense. This is equivalent to finding the minimum norm solutions to the compatible systems

$$A_kX = B_k \quad \text{and} \quad \bar{A}_kX = \bar{B}_k \tag{24}$$

where $B_k = A_kA_k^\dagger B$ and $\bar{B}_k = \bar{A}_k\bar{A}_k^\dagger\bar{B}$ are orthogonal projections of B and $\bar{B}$, respectively. From the results in §2, it remains to bound the subspace angle $\sin\theta_{LS}$:

$$\begin{aligned}\sin\theta_{LS} &= \text{dist}(\mathcal{N}([A_k\ B_k]), \mathcal{N}([\bar{A}_k\ \bar{B}_k])) \\ &= \|[A_k\ B_k]^\dagger[A_k\ B_k] - [\bar{A}_k\ \bar{B}_k]^\dagger[\bar{A}_k\ \bar{B}_k]\| \\ &\le \|[A_k\ B_k]^\dagger\|\,\|[\Delta A\ \Delta B]\| + t_{LS}\end{aligned}$$

where

$$t_{LS} \equiv \mu\,\|[A_k\ B_k]^\dagger\|\,\|[\bar{A}_k\ \bar{B}_k]^\dagger\|\,\|[A_k\ B_k] - [\bar{A}_k\ \bar{B}_k]\|\,\|[\Delta\bar{A}_k\ \bar{R}_k]\|$$

and $\bar{R}_k = \bar{B} - \bar{A}\,\bar{x}_{LS}$ is the residual. It remains to bound the individual terms. By the interlacing property of singular values [7, p.428] it follows

$$\|[A_k\ B_k]^\dagger\| \le \frac{1}{\sigma'_k} \qquad \text{and} \qquad \|[\bar{A}_k\ \bar{B}_k]^\dagger\| \le \frac{1}{\sigma'_k - \|\Delta A\|}.$$

If we define

$$P_{U'_k} = [u'_1, u'_2, \ldots, u'_k][u'_1, u'_2, \ldots, u'_k]^T \text{ and } P_{\bar{U}'_k} = [\bar{u}'_1, \bar{u}'_2, \ldots, \bar{u}'_k][\bar{u}'_1, \bar{u}'_2, \ldots, \bar{u}'_k]^T$$

then

$$\begin{aligned} \|[A_k\ B_k] - [\bar{A}_k\ \bar{B}_k]\| &\leq \|P_{U'_k}[A\ B] - P_{\bar{U}'_k}[\bar{A}\ \bar{B}]\| \\ &= \|(P_{U'_k} - P_{\bar{U}'_k})[A\ B] - P_{\bar{U}'_k}[\Delta A\ \Delta B]\| \\ &\leq \sin\beta_{LS}\|[A\ B]\| + \|[\Delta A\ \Delta B]\|. \end{aligned}$$

Here, $\sin\beta_{LS} \equiv \|P_{U'_k} - P_{\bar{U}'_k}\|$ and β_{LS} denotes the largest angle between $\mathcal{R}(A_k) = \mathcal{R}([A_k\ B_k])$ and $\mathcal{R}(\bar{A}_k) = \mathcal{R}([\bar{A}_k\ \bar{B}_k])$. Further, it follows from Wedin's [18] perturbation bounds for singular subspaces

$$\sin\beta_{LS} \leq \frac{\|\Delta A\|}{\sigma'_k - \sigma'_{k+1} - \|\Delta A\|}.$$

Define α_{LS} by

$$\alpha_{LS} = \mu\frac{(\sin\beta_{LS}\|[A\ B]\| + \|[\Delta A\ \Delta B]\|)\ \|[\Delta\bar{A}_k\ \bar{R}_k]\|}{\sigma'_k\,(\sigma'_k - \|\Delta A\|)}. \tag{25}$$

Then $t_{LS} \leq \alpha_{LS}$ and therefore

$$\sin\theta_{LS} \leq \frac{\|[\Delta A\ \Delta B]\|}{\sigma'_k} + \alpha_{LS}. \tag{26}$$

Under the reasonable circumstances $A_kX \approx B$ is not too incompatible and $\|[\Delta A\ \Delta B]\|$ is not too large, the first term dominates α_{LS}. If the upper bound is not too pessimistic, then $\sin\theta_{LS} \approx \frac{\|[\Delta A\ \Delta B]\|}{\sigma'_k}$.

Now we will find upper bounds for TLS. Let $[\check{A}\ \check{B}] = \sum_{i=1}^{k}\bar{\sigma}_i\bar{u}_i\bar{v}_i^T$ denote the nearest rank-k approximation to $[\bar{A}\ \bar{B}] = \sum_{i=1}^{n+d}\bar{\sigma}_i\bar{u}_i\bar{v}_i^T$. Then TLS finds the minimum norm solutions to the compatible systems

$$\hat{A}X = \hat{B} \quad \text{and} \quad \check{A}X = \check{B}. \tag{27}$$

In a similar manner as above,

$$\sin\theta_{TLS} = \operatorname{dist}(\mathcal{N}([\hat{A}\ \hat{B}]), \mathcal{N}([\check{A}\ \check{B}])) \tag{28}$$

$$= \|[\hat{A}\ \hat{B}]^\dagger[\hat{A}\ \hat{B}] - [\check{A}\ \check{B}]^\dagger[\check{A}\ \check{B}]\| \tag{29}$$

$$\leq \frac{\|[\Delta A\ \Delta B]\|}{\sigma_k} + t_{TLS}$$

where

$$t_{TLS} \equiv \mu\|[\hat{A}\ \hat{B}]^\dagger\|\ \|[\check{A}\ \check{B}]^\dagger\|\ \|[\hat{A}\ \hat{B}] - [\check{A}\ \check{B}]\|\ \|[\Delta\check{A}\ \Delta\check{B}]\|. \tag{30}$$

Note that

$$\|[\hat{A}\ \hat{B}]^\dagger\| \leq \frac{1}{\sigma_k} \qquad \text{and} \qquad \|[\check{A}\ \check{B}]^\dagger\| \leq \frac{1}{\sigma_k - \|[\Delta A\ \Delta B]\|}.$$

If we define

$$P_{U_k} = [u_1, u_2, \ldots, u_k][u_1, u_2, \ldots, u_k]^T \text{ and } P_{\bar{U}_k} = [\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_k][\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_k]^T$$

then

$$\begin{aligned} \|[\hat{A}\ \hat{B}] - [\check{A}\ \check{B}]\| &= \|(P_{U_k} - P_{\bar{U}_k})[A\ B] - P_{\bar{U}_k}[\triangle A\ \triangle B]\| \\ &\leq \sin\beta_{TLS}\|[A\ B]\| + \|[\triangle A\ \triangle B]\|, \end{aligned}$$

where $\sin\beta_{TLS} \equiv \|P_{U_k} - P_{\bar{U}_k}\|$ and β_{TLS} denotes the largest angle between $\mathcal{R}(\hat{A}) = \mathcal{R}([\hat{A}\ \hat{B}])$ and $\mathcal{R}(\check{A}) = \mathcal{R}([\check{A}\ \check{B}])$. Further, from the perturbation of singular subspaces in [16, 18], it follows

$$\sin\beta_{TLS} \leq \frac{\|[\triangle A\ \triangle B]\|}{\sigma_k - \sigma_{k+1} - \|[\triangle A\ \triangle B]\|}.$$

Define α_{TLS} by

$$\alpha_{TLS} = \mu \frac{(\sin\beta_{TLS}\|[A\ B]\| + \|[\triangle A\ \triangle B]\|)\ \|[\triangle \check{A}\ \ \triangle \check{B}]\|}{\sigma_k\,(\sigma_k - \|[\triangle A\ \triangle B]\|)}. \tag{31}$$

Then it follows $t_{TLS} \leq \alpha_{TLS}$ and

$$\sin\theta_{TLS} \leq \frac{\|[\triangle A\ \triangle B]\|}{\sigma_k} + \alpha_{TLS}. \tag{32}$$

Under reasonable circumstances as above, the first term dominates α_{TLS}. When the upper bound is not overly pessimistic then $\sin\theta_{TLS} \approx \frac{\|[\triangle A\ \triangle B]\|}{\sigma_k}$. Thus we have proved the following result.

THEOREM 3.2. *Let A and $[A\ B]$ have the SVD as in (2) and (5), respectively, and let $[\bar{A}\ \bar{B}] = [A\ B] + [\triangle A\ \triangle B]$. Let X_{LS}, $\bar{X}_{LS}$ and X_{TLS}, $\bar{X}_{TLS}$ denote the minimum norm solutions to (24) and (27), respectively, provided $\|[\triangle A\ \triangle B]\| < \sigma'_k - \sigma_{k+1}$. Then the following perturbation bounds hold:*

$$\begin{aligned} \sin\phi_{LS} \leq \quad \|X_{LS} - \bar{X}_{LS}\| &\leq \sin\phi_{LS}\sqrt{1 + \|X_{LS}\|^2}\,\sqrt{1 + \|\bar{X}_{LS}\|^2} \\ &\leq \left(\frac{\|[\triangle A\ \triangle B]\|}{\sigma'_k} + \alpha_{LS} + \epsilon_{LS}\right)\sqrt{1 + \|X_{LS}\|^2}\,\sqrt{1 + \|\bar{X}_{LS}\|^2} \end{aligned}$$

$$\begin{aligned} \sin\phi_{TLS} \leq \quad \|X_{TLS} - \bar{X}_{TLS}\| &\leq \sin\phi_{TLS}\sqrt{1 + \|X_{TLS}\|^2}\,\sqrt{1 + \|\bar{X}_{TLS}\|^2} \\ &\leq \left(\frac{\|[\triangle A\ \triangle B]\|}{\sigma_k} + \alpha_{TLS} + \epsilon_{TLS}\right)\sqrt{1 + \|X_{TLS}\|^2}\,\sqrt{1 + \|\bar{X}_{TLS}\|^2}. \end{aligned}$$

Neglecting ϵ_{LS} and α_{LS}, as well as ϵ_{TLS} and α_{TLS}, it appears the upper bound for TLS subspace angle is slightly smaller than TLS, since $\sigma'_k \leq \sigma_k$. Finally, σ_k can increase when the columns of B have an nontrivial orientation along the k^{th} left singular vector of A, namely, u'_k.

4. Bounds on $\|X_{LS} - X_{TLS}\|$. In this section we examine bounds on $\|X_{LS} - X_{TLS}\|$ using the general perturbation bound as a platform. Other bounds are given in [19]. As usual, we shall assume the general condition

$$\sigma'_k > \sigma_{k+1} \geq \cdots \geq \sigma_{n+d} \quad \text{for } k \leq n.$$

THEOREM 4.1. *Let A and $[A\ B]$ have the usual SVD. Let $R_k = B - AX_{LS}$, $X_{LS} = -C'\Gamma'^{-1}$, $X_{TLS} = -\hat{C}\hat{\Gamma}^{-1}$. If $\sigma_{k+1} < \sigma'_k$ then*

$$(33)\quad \sin\phi \le \|X_{LS} - X_{TLS}\| \le \left(\mu \frac{\sigma_{k+1}(2\sigma_{k+1} + \|R_k\|)}{{\sigma'_k}^2} + \epsilon\right)\sqrt{1+\|X_{LS}\|^2}\sqrt{1+\|X_{TLS}\|^2}$$

where ϕ is the subspace angle between $\mathcal{R}([C'^T\ \Gamma'^T]^T)$ and $\mathcal{R}([\hat{C}^T\ \hat{\Gamma}^T]^T)$ and $\mu = (1+\sqrt{5})/2$.

Proof: Using $\sin\phi \le \sin\theta + \epsilon$, Theorems 2.2 and 2.5, we can set $\|[\Delta A\ \Delta b]\| = 0$, and then we need only bound $\sin\theta$:

$$\begin{aligned} \sin\theta &= \|[\hat{A}\ \hat{B}]^\dagger[\hat{A}\ \hat{B}] - [A_k\ B_k]^\dagger[A_k\ B_k]\| \\ &\le \mu\|[\hat{A}\ \hat{B}]^\dagger\|\,\|[A_k\ B_k]^\dagger\|\,\|[\hat{A}\ \hat{B}] - [A_k\ B_k]\|\,\|[\Delta\hat{A}\ \Delta\hat{B}]\| \\ &\le \mu\frac{1}{\sigma_k}\frac{1}{\sigma'_k}(2\sigma_{k+1} + \|R_k\|)\sigma_{k+1} \\ &\le \frac{\mu}{{\sigma'_k}^2}(2\sigma_{k+1} + \|R_k\|)\sigma_{k+1}. \end{aligned}$$

Note that if (1) is compatible and $\text{rank}(A) = k$, then $X_{LS} = X_{TLS}$, as expected from [14]. The proof of Theorem 4.1 shows that when A has a well-determined gap in its singular value spectrum and $A_kX \approx B$ is not too incompatible, then the subspace angle between the LS and TLS approximate nullspace is $\mathcal{O}(\sigma_{k+1}/\sigma'_k)^2$ (whenever ϵ is sufficiently small). Numerical experiments (cf. Table 5) confirm this observation. By Theorem 4.1, we conclude provided the LS and TLS problems are not too ill-conditioned (i.e., $\sqrt{1+\|X_{LS}\|^2}$ and $\sqrt{1+\|X_{TLS}\|^2}$ is not too large), this ratio then plays an influential role in determining the similarities and differences by *estimating* the angle between the LS and TLS approximate nullspaces. Table 5 illustrates $\sin\phi \approx \mathcal{O}(\sigma_{k+1}/\sigma'_k)^2$ and summarizes the results of numerical simulations which are relevant to Theorem 4.1.

5. Numerical Simulations. In this section we present the results of some computer simulations to illustrate our theory. In the following Matlab computer simulations we considered the problem of resolving sinusoids with closely spaced frequencies in a noisy environment using the forward linear prediction (FLP) model in root form.

We estimate the FLP coefficients $\{\bar{x}_i\}$ from the noisy problem

$$\begin{pmatrix} \bar{a}_L & \bar{a}_{L-1} & \cdots & \bar{a}_1 \\ \bar{a}_{L+1} & \bar{a}_L & \cdots & \bar{a}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{N-1} & \bar{a}_{N-2} & \cdots & \bar{a}_{N-L} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_L \end{pmatrix} \approx \begin{pmatrix} \bar{a}_{L+1} \\ \bar{a}_{L+2} \\ \vdots \\ \bar{a}_N \end{pmatrix},$$

which is denoted

$$(34)\qquad \bar{A}x \approx \bar{b}.$$

Here, $[\bar{A}\ \bar{b}] = [A\ b] + [\Delta A\ \Delta b]$, $Ax = b$ is compatible, and $\text{rank}(A) = p$, where p is the number of complex exponentials. The dimensions of the problem are given by $\bar{A}$ is $(N-L)\times L$, x is $L\times 1$, $\bar{b}$ is $(N-L)\times 1$, and $(N-L) > L > p$. Finally, N is the number of measured observations and L is the prediction order.

In the following two cases $p = 4$, the exact solution x_0 satisfies $\|x_0\| = 1.1214$, and the frequencies are spaced according to $f_1 = 0.450$ Hz and $f_2 = 0.459$ Hz.

- In Case 1 we set $N = 55$, $L = 14$. The singular value spectrum of A is $\sigma(A) = \{0.19250,\ 0.14654,\ 0.16240\cdot 10^{-1},\ 0.49645\cdot 10^{-2},\ 0,\ \ldots,\ 0\}$.

- In Case 2 we set $N = 65$, $L = 14$. The singular value spectrum of A is $\sigma(A) = \{0.19284,\ 0.14732,\ 0.23800 \cdot 10^{-1},\ 0.73616 \cdot 10^{-2},\ 0,\ \ldots,\ 0\}$.

In both cases we perturbed the compatible system $Ax = b$ by structure-preserving noise $[\Delta A\ \Delta b]$ from the normal distribution with mean zero. The standard deviation of the noise, σ, ranged from $7 \cdot \times 10^{-5}$ to $1 \cdot 10^{-3}$. Table 1 summarizes the LS and TLS results for Case 1 and Table 2 summarizes the results for the corresponding subspace angles. Table 3 summarizes the results for Case 2 and Table 4 summarizes the results for the corresponding subspace angles. The values in these tables represent the average of the outcomes of 50 trials. The results of each trial are captured in the histograms in Figure 1.

In Tables 1 and 3 we see that the lower and upper bounds provide realistic estimates of the error. As suggested by Theorem 3.1, our numerical results indicate that the method with the smaller subspace angle produces a more accurate solution. These simulations illustrate the TLS subspaces usually filter more noise than the LS subspaces, and consequently, the TLS solutions provide more accurate estimates than LS in the perturbation of zero-residual problems. In addition, as suggested by Theorem 3.1, Tables 1 and 3 illustrate that a larger k^{th} singular value ($0.49645 \cdot 10^{-2}$ vs. $0.73616 \cdot 10^{-2}$ for LS and $0.64353 \cdot 10^{-2}$ vs. $0.94786 \cdot 10^{-2}$ for TLS) produces less sensitive subspaces, hence better parameter estimates.

In Tables 2 and 4 we see that $\sin\phi_M$ is realistically appraised by $\sin\theta_M$ and that mostly $\epsilon_M \leq \sin\theta_M$ (M=LS or TLS).

Now we discuss some computer simulations relevant to Theorem 4.1. In the following Matlab computer simulations, we chose $A \in \Re^{30\times 10}$ with the singular value spectrum satisfying numerical rank(A) < rank(A) as follows:

$$\sigma(A) = \{\,1,\ 0.5,\ 0.2,\ 0.1,\ 0.05,\ 0.03,\ 0.01,\ \sigma_8',\ \sigma_9',\ \sigma_{10}'\,\} \tag{35}$$

where σ_8', σ_9', and σ_{10}' vary as follows:

$$\begin{aligned}
\text{Case 3: } \{\,\sigma_8', \sigma_9', \sigma_{10}'\,\} &= \{\,10^{-9}, 10^{-10}, 10^{-11}\,\} \\
\text{Case 4: } \{\,\sigma_8', \sigma_9', \sigma_{10}'\,\} &= \{\,10^{-6}, 10^{-7}, 10^{-8}\,\} \\
\text{Case 5: } \{\,\sigma_8', \sigma_9', \sigma_{10}'\,\} &= \{\,10^{-4}, 10^{-5}, 10^{-6}\,\}.
\end{aligned}$$

For Cases $3-5$ we chose a random matrix $X_i \in \Re^{10\times 3}$ from the standard normal distribution and set $B_i = AX_i$ ($2.5 \leq \|X_i\| \leq 4$ and $1.4 \leq \|B_i\| \leq 2.0$). Then we perturbed the system $AX_i = B_i$ ($i = 3,4,5$) by noise $[\Delta A\ \Delta B]$ from the normal distribution with mean 0. In Case 3, the standard deviation of the noise ranged from 10^{-11} to 10^{-4}. In Case 4, it ranged from 10^{-8} to 10^{-4}. In Case 5, from 10^{-6} to 10^{-4}. For fixed i, X_{LS} and X_{TLS} denote the truncated LS and TLS solution, resp., to $AX = B_i$ and $\bar{X}_{LS}$ and $\bar{X}_{TLS}$ denote the perturbed solutions. We remark that although $AX = B_i$ is compatible, the corresponding LS or TLS residual is usually nonzero due to truncation.

Table 5 summarizes the results for $\|X_{LS} - X_{TLS}\|$ and its bounds, $\sin\theta$, and the scalar $\sigma_8/\sigma_7')^2$. All values represent the average of the outcomes for 100 trials. The table shows $(\sigma_8/\sigma_7')^2$ estimates the $\|X_{LS} - X_{TLS}\|$ quite well.

6. Conclusions. In this paper new perturbation theory is presented for orthogonal projection methods with applications to LS and TLS. A model for orthogonal projection methods is established (§2) to "solve" the overdetermined system of linear equations $AX \approx B$. It is shown that the minimum norm solution X_M is competely determined by the kernel of the lower rank approximation to $[A\ B]$. Also, interesting relationships between the solution and an orthonormal basis for the kernel are proven. If M$\equiv$TLS, then the model specializes to well known TLS model.

Lower and upper perturbation bounds for the solution using method M are presented (§2). The bounds are in terms of the subspace angle between approximate nullspaces and the norms of the solutions. This shows that method M makes only assumptions about the noise distribution in the kernel of the lower rank approximation to $[A\ B]$, and that various perturbations may produce the same solution. Then it is shown how the subspace angle can be bounded (§2) in terms of the fittings, corrections, and perturbation. This leads to new perturbation bounds for LS and TLS (§3). Numerical simulations show the LS and TLS bounds involving the subspace angle are quite good.

The analysis shows that usually the TLS upper bound is slightly smaller than the corresponding LS perturbation bound in the presence of noise in all the data (§3), and hence the TLS subspace is less sensitive (smaller subspace angle) to noise than the LS subspace. This is confirmed by numerical simulations in the perturbation of rank-deficient compatible systems (§5). This explains the superiority of TLS over LS observed in the literature.

Furthermore, the general perturbation bounds also permit a comparison between any two competing orthogonal projection methods. In particular, the LS and TLS solutions are compared. The bounds identify the subspace angle as a key factor in determining the similarities and differences between the two solutions, and the numerical results show the quantity $(\sigma_{k+1}/\sigma'_k)^2$ is shown to be closely related to $\|X_{LS} - X_{TLS}\|$ (§4).

Acknowledgement. The authors wish to thank Sabine Van Huffel, A. van der Sluis, and Musheng Wei for their helpful suggestions. The first author wishes to thank the Department of Mathematics at UCLA for use of their computing facilities.

REFERENCES

[1] Ezio Biglieri and Kung Yao, *Some Properties of Singular Value Decomposition and their Applications to Digital Signal Processing*, Signal Processing, 18, (1989), pp. 277-289.

[2] A. Björck, *Least Squares Methods*, in *Handbook of Numerical Analysis, Volume I: Finite Difference Methods-Solution of Equations in R^n*, edited by P. G. Ciarlet and J. L. Lions, Elsevier/North Holland, 1990.

[3] T.F. Chan and P.C. Hansen, *Some Applications of the Rank Revealing QR Factorization*, SIAM J. Sci. Stat. Comput., 13, (1992), pp. 727-741.

[4] Ricardo D. Fierro and James R. Bunch, *Orthogonal Projection and Total Least Squares*, Numer. Lin. Alg. Applic., accepted.

[5] G.H. Golub and V. Klema and G.W. Stewart, *Rank Degeneracy and Least Squares Problems*, Technical Report TR-456, Computer Science Department, University of Maryland, 1976.

[6] 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.

[7] G.H. Golub and C.F. Van Loan, *Matrix Computations (Second Edition)*, John Hopkins University Press , 1989.

[8] P.C. Hansen, *The Truncated SVD as a Method for Regularization*, BIT, 27, (1987), pp. 534-553.

[9] C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*, Prentice Hall, 1974.

[10] C.C. Paige and M.A. Saunders, *Towards a Generalized Singular Value Decomposition*, SIAM J. Numer. Anal., 18, (1981), pp. 398-405.

[11] G.W. Stewart, *An Updating Algorithm for Subspace Tracking*, Technical Report UMIACS-TR 90-86, CS-TR 2494, Computer Science Department, University of Maryland, MD, 1990.

[12] Petre Stoica and Torsten Soderstrom, *On Spectral and Root Forms of Sinusoidal Frequency Estimators* , IEEE Proceedings, (1991), pp. 3257-3260.

[13] A. van der Sluis and G.W. Veltkamp, *Restoring Rank and Consistency by Orthogonal Projection*, Lin. Alg. Appl., 28, (1979), pp. 257-278.

[14] S. Van Huffel and J. Vandewalle, *Algebraic Relationships Between Classical Regression and Total Least Squares Estimation*, Lin. Alg. Appl., 93, (1987), pp. 149-162.

[15] S. Van Huffel and J. Vandewalle, *Algebraic Connections Between the Least Squares and Total Least Squares Problems*, Numer. Math., 55, (1989), pp. 431-449.

[16] S. Van Huffel and J. Vandewalle, The Total Least Squares Problem: Computational Aspects and Analysis, SIAM, 1990.

[17] S. Van Huffel and H. Zha, *An Efficient Total Least Squares Algorithm Based on a Rank-Revealing Two-Sided Orthogonal Decomposition*, Technical Report ESAT-SISTA 1991-08, Department of Electrical Engineering, Katholieke Universisteit Leuven, Belgium, 1991.

[18] Per Åke Wedin, *Perturbation Theory for Pseudoinverses*, BIT, 13, (1973), pp. 217-232.

[19] Musheng Wei, *Algebraic Relations Between the Total Least Squares and Least Squares Problems with More than One Solution*, Numer. Math., 62, (1992), pp. 123-148.

[20] Kai-bor Yu and Md. Anisur Rahman, *Total least squares approach for frequency estimation using linear prediction*, IEEE Transactions on Acoustics, Speech and Signal Processing, 35, (1987), pp. 277-289.

[21] M.D. Zoltowski, *Generalized Miminum Norm and Constrained Total Least Squares with Applications to Signal Processing*, Advanced Architectures and Algorithms for Signal Processing III, SPIE, 975, (1988), pp. 78-85.

σ	$\|f_0 - f_{LS}\|$	$\sin\phi_{LS}$	$\|x_0 - \bar{x}_{LS}\|$	$\sin\phi_{LS}\sqrt{\cdot}\sqrt{\cdot}$
7.0e-05	4.0441e-04	1.1852e-02	1.9107e-02	2.6658e-02
9.0e-05	6.2276e-04	1.7439e-02	2.8477e-02	3.9141e-02
1.0e-04	7.1727e-04	1.9075e-02	3.1535e-02	4.2703e-02
3.0e-04	2.9938e-03	9.4030e-02	1.6094e-01	1.9953e-01
5.0e-04	8.2532e-02	2.2696e-01	3.7142e-01	4.4328e-01
7.0e-04	2.3703e-01	4.0881e-01	6.3515e-01	7.4136e-01
9.0e-04	2.0758e-01	5.2861e-01	8.0701e-01	9.1886e-01
1.0e-03	2.5522e-01	5.6893e-01	8.6304e-01	9.7877e-01
σ	$\|f_0 - f_{TLS}\|$	$\sin\phi_{TLS}$	$\|x_0 - \bar{x}_{TLS}\|$	$\sin\phi_{TLS}\sqrt{\cdot}\sqrt{\cdot}$
7.0e-05	3.5291e-04	1.1169e-02	1.7681e-02	2.5221e-02
9.0e-05	5.7031e-04	1.6297e-02	2.6136e-02	3.6823e-02
1.0e-04	6.0568e-04	1.6928e-02	2.7049e-02	3.8204e-02
3.0e-04	1.7934e-03	5.8442e-02	9.2296e-02	1.3185e-01
5.0e-04	2.8366e-03	1.1965e-01	1.8967e-01	2.7202e-01
7.0e-04	4.3220e-02	2.3929e-01	3.7800e-01	5.3751e-01
9.0e-04	1.8734e-01	4.0433e-01	6.2839e-01	8.3911e-01
1.0e-03	3.0808e-01	4.3969e-01	6.8327e-01	9.0609e-01

TABLE 1

Case 1: Comparison of truncated LS and truncated TLS bounds under perturbations of the compatible system $Ax = b$ by noise $[\Delta A\ \Delta b]$ from the normal distribution with mean 0 and standard deviation σ. The frequencies are $f_1 = 0.450$ Hz and $f_2 = 0.459$ Hz, and $N = 55$ and $L = 14$.

σ	$\sin\phi_{LS}$	$\sin\theta_{LS} + \epsilon_{LS}$	$\sin\theta_{LS}$
7.0e-05	1.1852e-02	2.0062e-02	1.3971e-02
9.0e-05	1.7439e-02	2.8956e-02	2.0255e-02
1.0e-04	1.9075e-02	3.1608e-02	2.2269e-02
3.0e-04	9.4030e-02	1.5231e-01	1.0588e-01
5.0e-04	2.2696e-01	3.9331e-01	2.6826e-01
7.0e-04	4.0881e-01	7.7458e-01	5.3256e-01
9.0e-04	5.2861e-01	1.0965e+00	7.3413e-01
1.0e-03	5.6893e-01	1.2082e+00	8.0349e-01
σ	$\sin\phi_{TLS}$	$\sin\theta_{TLS} + \epsilon_{TLS}$	$\sin\theta_{TLS}$
7.0e-05	1.1169e-02	1.8890e-02	1.3293e-02
9.0e-05	1.6297e-02	2.8381e-02	1.9150e-02
1.0e-04	1.6928e-02	2.9531e-02	2.0149e-02
3.0e-04	5.8442e-02	1.0246e-01	6.9508e-02
5.0e-04	1.1965e-01	2.0429e-01	1.4330e-01
7.0e-04	2.3929e-01	4.1235e-01	2.9021e-01
9.0e-04	4.0433e-01	7.6255e-01	5.3041e-01
1.0e-03	4.3969e-01	8.4100e-01	5.7799e-01

TABLE 2

Case 1: Comparison of $\sin\phi_{LS}$ and $\sin\theta_{LS}$, as well as $\sin\phi_{TLS}$ and $\sin\theta_{TLS}$. Here, the frequncies are $f_1 = 0.450$ Hz and $f_2 = 0.459$ Hz, and $N = 55$ and $L = 14$.

σ	$\|f_0 - f_{LS}\|$	$\sin\phi_{LS}$	$\|x_0 - \bar{x}_{LS}\|$	$\sin\phi_{LS}\sqrt{\cdot}\sqrt{\cdot}$
7.0e-05	3.3610e-04	6.6645e-03	1.0884e-02	1.5012e-02
9.0e-05	4.5089e-04	8.5048e-03	1.3877e-02	1.9151e-02
1.0e-04	5.1025e-04	1.0338e-02	1.6925e-02	2.3237e-02
3.0e-04	2.0950e-03	4.9285e-02	8.6829e-02	1.0746e-01
5.0e-04	3.4182e-03	1.2037e-01	2.0965e-01	2.4994e-01
7.0e-04	7.2137e-02	2.2184e-01	3.6528e-01	4.3398e-01
9.0e-04	1.6100e-01	3.3563e-01	5.3320e-01	6.1917e-01
1.0e-03	2.2166e-01	4.1923e-01	6.5215e-01	7.5308e-01
σ	$\|f_0 - f_{TLS}\|$	$\sin\phi_{TLS}$	$\|x_0 - \bar{x}_{TLS}\|$	$\sin\phi_{TLS}\sqrt{\cdot}\sqrt{\cdot}$
7.0e-05	3.0564e-04	6.1604e-03	9.7787e-03	1.3907e-02
9.0e-05	4.2490e-04	8.1620e-03	1.3040e-02	1.8444e-02
1.0e-04	4.7421e-04	9.4265e-03	1.4995e-02	2.1289e-02
3.0e-04	1.3079e-03	2.9750e-02	4.7237e-02	6.7263e-02
5.0e-04	1.8015e-03	5.4272e-02	8.5342e-02	1.2223e-01
7.0e-04	2.7937e-03	1.0471e-01	1.6574e-01	2.3871e-01
9.0e-04	1.3215e-02	1.6933e-01	2.6968e-01	3.8984e-01
1.0e-03	5.8285e-02	2.2887e-01	3.5628e-01	4.8818e-01

TABLE 3

Case 2: Comparison of truncated LS and truncated TLS bounds under perturbations of the compatible system $Ax = b$ by noise $[\triangle A\ \triangle b]$ from the normal distribution with mean 0 and standard deviation σ. The frequncies are $f_1 =$ 0.450 Hz and $f_2 =$ 0.459 Hz, and $N = 65$ and $L = 14$.

σ	$\sin\phi_{LS}$	$\sin\theta_{LS} + \epsilon_{LS}$	$\sin\theta_{LS}$
7.0e-05	6.6645e-03	1.2255e-02	8.2135e-03
9.0e-05	8.5048e-03	1.5628e-02	1.0332e-02
1.0e-04	1.0338e-02	1.9018e-02	1.2714e-02
3.0e-04	4.9285e-02	8.4540e-02	5.5317e-02
5.0e-04	1.2037e-01	1.9826e-01	1.3209e-01
7.0e-04	2.2184e-01	3.8278e-01	2.5948e-01
9.0e-04	3.3563e-01	6.0650e-01	4.1709e-01
1.0e-03	4.1923e-01	7.8111e-01	5.4511e-01
σ	$\sin\phi_{TLS}$	$\sin\theta_{TLS} + \epsilon_{TLS}$	$\sin\theta_{TLS}$
7.0e-05	6.1604e-03	1.1493e-02	7.7769e-03
9.0e-05	8.1620e-03	1.4365e-02	1.0075e-02
1.0e-04	9.4265e-03	1.7451e-02	1.1918e-02
3.0e-04	2.9750e-02	5.3809e-02	3.6994e-02
5.0e-04	5.4272e-02	9.7650e-02	6.7033e-02
7.0e-04	1.0471e-01	1.7926e-01	1.2737e-01
9.0e-04	1.6933e-01	2.8606e-01	2.0268e-01
1.0e-03	2.2887e-01	4.2287e-01	2.9795e-01

TABLE 4

Case 2: Comparison of $\sin\phi_{LS}$ and $\sin\theta_{LS}$, as well as $\sin\phi_{TLS}$ and $\sin\theta_{TLS}$. Here, the frequncies are $f_1 =$ 0.450 Hz and $f_2 =$ 0.459 Hz, and $N = 65$ and $L = 14$.

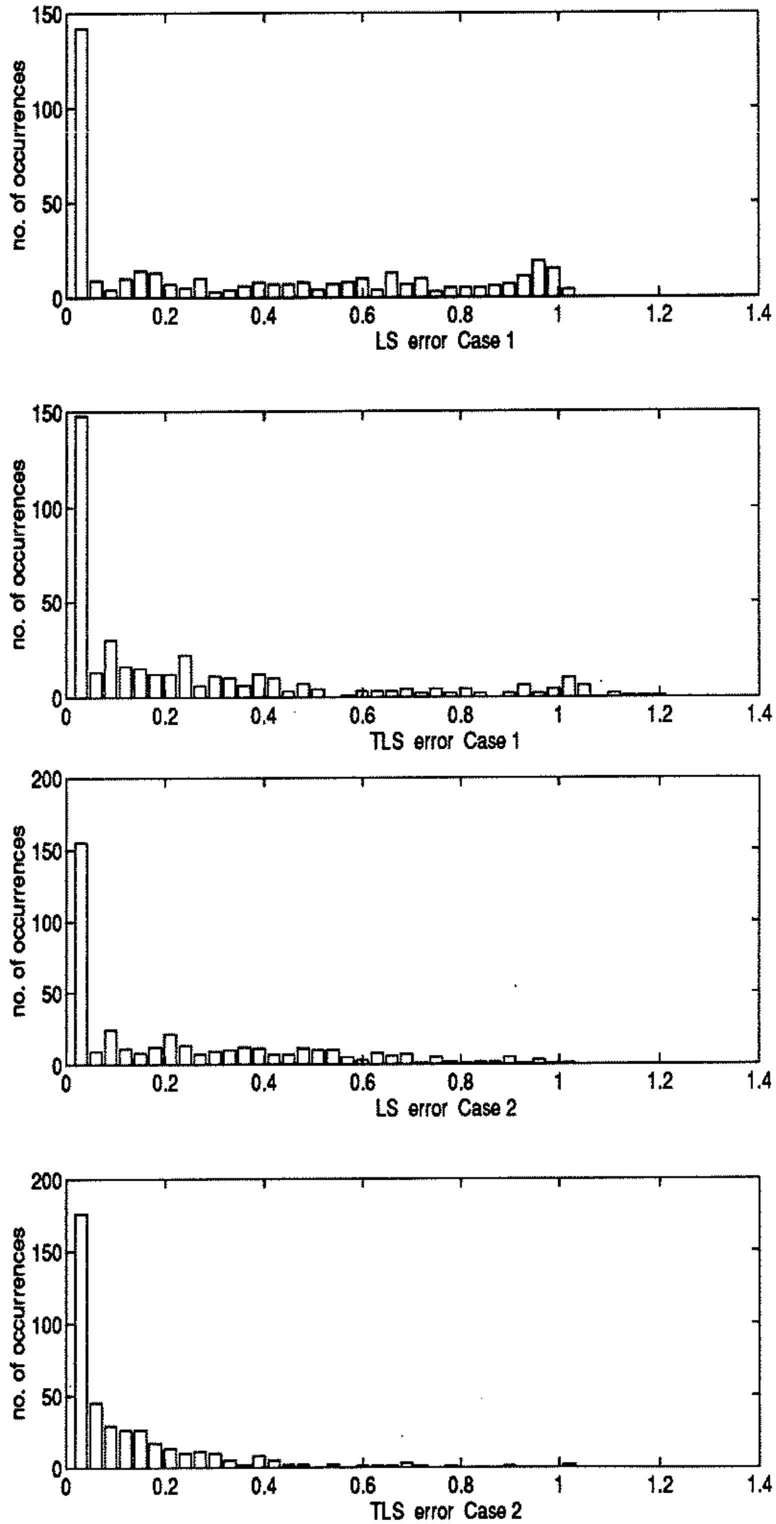

FIG. 1. *The frequency histograms of the truncated LS and TLS solutions to the noisy problems.*

i	σ	$\sin\phi$	$\|X_{LS} - X_{TLS}\|$	$\sin\phi\sqrt{\cdot}\sqrt{\cdot}$	$\sin\theta$	$(\sigma_8/\sigma_7')^2$
3	1.0e-11	4.9965e-15	1.1635e-14	3.0141e-14	4.2348e-15	2.3475e-14
	1.0e-10	5.2029e-15	1.3177e-14	3.1386e-14	4.4589e-15	2.6461e-14
	1.0e-09	1.2469e-13	4.3205e-13	7.5217e-13	1.1650e-13	4.4861e-13
	1.0e-08	1.2552e-11	4.4770e-11	7.5720e-11	1.1811e-11	4.2552e-11
	1.0e-07	1.2426e-09	4.3743e-09	7.4958e-09	1.1718e-09	4.2143e-09
	1.0e-06	1.2045e-07	4.2580e-07	7.2657e-07	1.1364e-07	4.2971e-07
	1.0e-05	1.2041e-05	4.3296e-05	7.2658e-05	1.1373e-05	4.2587e-05
	1.0e-04	1.1921e-03	4.2041e-03	7.1849e-03	1.1240e-03	4.3035e-03
4	1.0e-08	1.6488e-09	3.4774e-09	2.1827e-08	1.9087e-09	1.1457e-08
	1.0e-07	2.2968e-09	5.5373e-09	3.0405e-08	2.4920e-09	1.3999e-08
	1.0e-06	1.2709e-07	3.4808e-07	1.6823e-06	1.2914e-07	4.4896e-07
	1.0e-05	1.2274e-05	3.3037e-05	1.6248e-04	1.2474e-05	4.1894e-05
	1.0e-04	1.1695e-03	3.1715e-03	1.5487e-02	1.1888e-03	4.3070e-03
5	1.0e-06	2.1303e-06	7.2403e-06	2.6030e-05	2.1508e-06	1.1101e-04
	1.0e-05	1.1459e-05	4.5995e-05	1.4001e-04	1.2131e-05	1.3532e-04
	1.0e-04	1.0983e-03	4.4583e-03	1.3399e-02	1.1244e-03	4.2291e-03

TABLE 5

Comparison of $(\sigma_8/\sigma_7')^2$, $\sin\phi$ *and* $\sin\theta$ *for cases 3-5, where* $A \in \Re^{30\times 10}$ *and* $B_i \in \Re^{30\times 3}$. *The numerical rank of* A *is 7 in all cases. The simulation is described in* §5.