Katholieke Universiteit Leuven

Departement Elektrotechniek

ESAT-SISTA/TR 1996-58

Structured Total Least Squares Problems: Formulations, Algorithms and Applications*

Philippe Lemmerling, Sabine Van Huffel, and Bart De Moor†

September 1996
Published in Proceedings of 2nd International Workshop on TLS and Errors in Variables Modeling

*This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/lemmerli/reports/int9658.ps.Z*

†ESAT - Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven (Heverlee), Belgium, tel 32/16/32 17 96, fax 32/16/32 19 86, email: *philippe.lemmerling@esat.kuleuven.ac.be, sabine.vanhuffel@esat.kuleuven.ac.be, bart.demoor@esat.kuleuven.ac.be.*
This work is supported by the Belgian Programme on Interuniversity Poles of Attraction (IUAP-nr.17 and IUAP nr. 50), initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, by the EU Programme "Human Capital and Mobility, Networks", project CHRX-CT94–432, and by a Concerted Research Action (GOA) project of the Flemish Community, entitled "Model-based Information Processing Systems". The first author is a Ph.D. student funded by the IWT (Flemish Institute for Support of Scientific-Technological Research in Industry). The second and third authors are Research Associates with the Belgian N.F.W.O. (National Fund for Scientific Research).

Abstract

The Total Least Squares (TLS) approach is a popular method in data fitting problems where we have to solve an overdetermined system of linear equations $Ax \approx b$. When the noise on the entries of $[A \;\; b]$ is i.i.d. Gaussian white noise of equal variance, TLS yields a Maximum Likelihood (ML) estimate for x, ΔA and Δb such that $(A + \Delta A)x = b + \Delta b$. The latter statistical condition is not satisfied in many signal processing applications where we have to deal with structured (Hankel, Toeplitz, sparse matrices,...) data matrices $[A \;\; b]$ with i.i.d. Gausssian white noise of equal variance on the *different* entries. A necessary condition for ML estimates in that case is that the structure of $[\Delta A \;\; \Delta b]$ equals that of $[A \;\; b]$. Therefore several formulations and corresponding solution methods have been developed, that allow to fix the structure of $[\Delta A \;\; \Delta b]$, while minimizing $\|[\Delta A \;\; \Delta b]\|_F$.
We compare the formulation of three approaches: the Structured Total Least Norm (STLN) method, the Structured Total Least Squares (STLS) method and the Constrained Total Least Squares (CTLS) method. Although these three approaches have a different formulation and associated solution methods, we can show that they all solve the same problem and can be interpreted in an algebraic way as a pure optimization approach, a Singular Value Decomposition (SVD) and an Eigenvalue Decomposition (EVD) approach. We discuss the properties (convergence rate and accuracy) of the algorithms for solving the different formulations. Finally, we illustrate these properties, as well as the efficiency of the applied algorithms, on a simple modeling example.

1 Introduction

The Total Least Squares (TLS) method is a frequently used method in data fitting problems of the form

$$Ax \approx b, \quad A \in \mathcal{R}^{m \times n}, \quad b \in \mathcal{R}^{m \times 1}, \quad m > n. \tag{1}$$

Since the entries of A and b usually are measurements obtained from an experimental setup, they consist of 2 parts: the true unknown part (A_0 and b_0) and another part which represents noise (ΔA and Δb):

$$\begin{aligned} A &= A_0 + \Delta A \\ b &= b_0 + \Delta b. \end{aligned}$$

In most applications one is interested in finding a Maximum Likelihood (ML) estimate of the parameter vector x and the corrections ΔA and Δb. From [13] it is well known that the TLS method applied to (1) only yields ML estimates when the noise on the entries of $[A \;\; b]$ is i.i.d. Gaussian white noise of equal variance. However the latter statistical condition does not hold for numerous applications where the matrix $[A \;\; b]$ has some structure (Hankel matrix, Toeplitz matrix, sparse matrix, ...).
Intuitively, it is easy to understand that for the case in which the noise on the different entries of $[A \;\; b]$ is i.i.d. Gaussian white noise of equal variance, a necessary condition for a ML estimate is that $[\Delta A \;\; \Delta b]$ should have the same structure as $[A \;\; b]$ (a proof can be found in [2]). Therefore, several approaches have been developed for solving (1) while preserving the structure of the original extended data matrix $[A \;\; b]$:

$$\begin{aligned} &\min_{\Delta A, \Delta b, x} \|[\Delta A \;\; \Delta b]\|_F \\ &\text{such that } \; (A + \Delta A)x = (b + \Delta b) \; \text{ and} \\ &[\Delta A \;\; \Delta b] \text{ has the same structure as } [A \;\; b]. \end{aligned} \tag{2}$$

This looks like a simple extension of the well-known TLS approach (there is only one supplementary constraint). To illustrate that this is not the case, we reformulate the ordinary TLS and the structured TLS formulation (2) into unconstrained optimization problems that are only a function of the vector $y \in \mathcal{R}^{(n+1)\times 1}$, with y such that $[A+\Delta A \quad b+\Delta b]y = 0$. Note that the replacement of $[x^T \; -1]^T$ by y is permitted in generic cases (the nongeneric cases are those in which the $(n+1)$th entry of y is 0 [13]). Since it is obvious that if y is a solution, then also $\alpha y, \alpha \neq 0$, is a solution, we can restrict ourselves to the y vectors on the unit sphere. As an example we take

$$A = \begin{bmatrix} 0.2190 & 0.0470 \\ 0.9347 & 0.2190 \\ 0.3835 & 0.9347 \end{bmatrix}, b = \begin{bmatrix} 0.6789 \\ 0.0470 \\ 0.2190 \end{bmatrix}.$$

By using spherical coordinates θ and ϕ (see Figure 1), we get Figure 2 for TLS and Figure 3 for the structured TLS case. For reasons of symmetry we only have to examine one quadrant of the unit sphere (e.g. the upper left part of the contour plot). In Figure 2 we see that TLS yields one unique minimum (see cross in rectangular area on contour plot). In Figure 3 we see why the additional constraint is not just a simple extension of TLS: we have multiple minima (see crosses in the rectangular areas on contour plot).

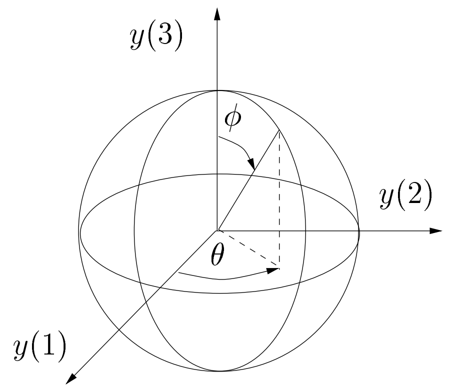

FIG. 1. Identification of a point with Euclidean coordinates $(y(1), y(2), y(3))$ on the unit sphere by spherical coordinates (ϕ, θ)

The paper is structured as follows. Section 2 briefly describes the formulation of the three approaches: the Constrained Total Least Squares (CTLS) approach, the Structured Total Least Norm (STLN) approach and the Structured Total Least Squares (STLS) approach. We also discuss possible extensions of these formulations towards other norms and multiple right-hand side problems (where we have $B \in \mathcal{R}^{m\times d}$ instead of $b \in \mathcal{R}^{m\times 1}$). Next we give solution methods for the different problem formulations. Section 4 gives an interpretation of the solution methods corresponding to the different problem formulations.
The findings of the previous sections are confirmed by a numerical example in Section 5. Finally, we formulate our conclusions and indicate ways for further research.
In the remainder of the paper we adopt a Matlab-like notation: $v(i)$ indicates the ith element of vector v, $A(i,j)$ indicates the entry of A on row i and in column j and $A(:,i)$ denotes the ith column of A. We will use the following shorthand notations for the extended data matrix and its corresponding correction matrix: $S \equiv [A \;\; b]$ and $\Delta S \equiv [\Delta A \;\; \Delta b]$.

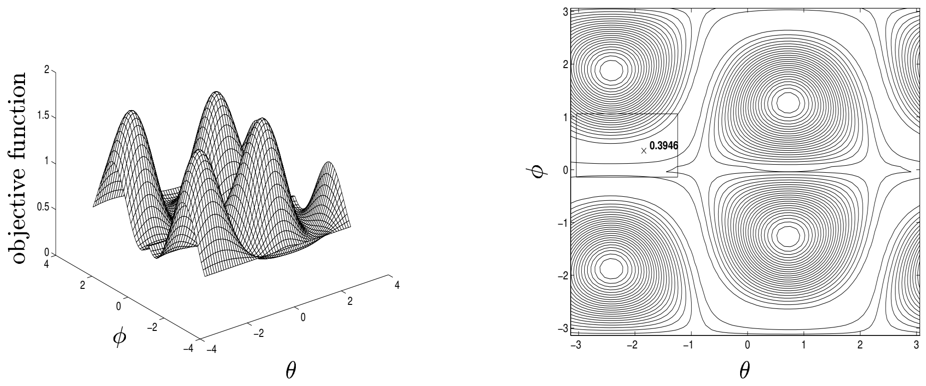

FIG. 2. The left figure shows a mesh plot of the objective function that we have to minimize for solving the TLS problem. The right figure shows the corresponding contour plot. The coordinates θ and ϕ are spherical coordinates. The cross indicates the minimum in the upper left quadrant. The number is the function value at that point.

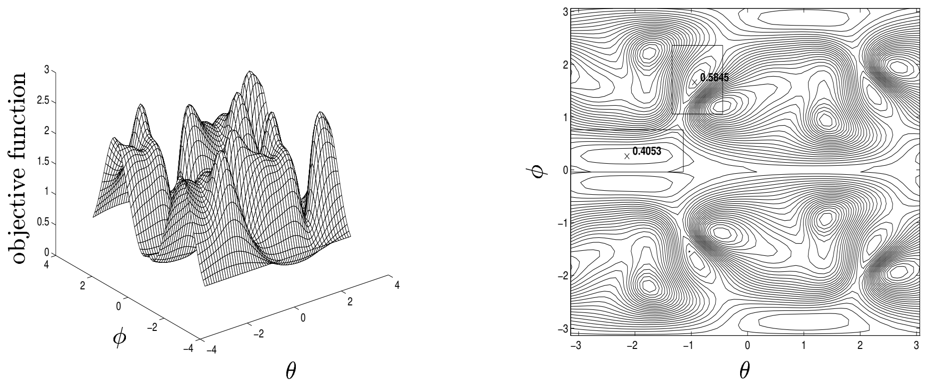

FIG. 3. The left figure shows a mesh plot of the objective function that we have to minimize for solving the structured TLS problem. The right figure shows the corresponding contour plot. The coordinates θ and ϕ are spherical coordinates. The crosses indicate the minima in the upper left quadrant. The accompanying numbers are the corresponding function values.

2 Formulations

In this section we describe three structured TLS formulations. To fix thoughts, we illustrate each of the formulations by means of the following Hankel matrix example:

$$S = \begin{bmatrix} 5 & 6 & 7 \\ 6 & 7 & 8 \end{bmatrix}^T, \Delta S = \begin{bmatrix} \tilde{\alpha} & \tilde{\beta} & \tilde{\gamma} \\ \tilde{\beta} & \tilde{\gamma} & \tilde{\delta} \end{bmatrix}^T. \tag{3}$$

An important idea behind all these formulations is that they try to represent the data by a minimal number of parameters while the original structure is memorized in matrices of a specific format.

2.1 CTLS

The CTLS formulation uses a so-called noise vector $f \in \mathcal{R}^{k\times 1}$ to represent the different elements of ΔS. The structure is captured by matrices $F_i \in \mathcal{R}^{m\times k}$ such that:

$$\Delta S(:,i) = F_i f, \;\; i = 1,\ldots,n+1.$$

For our Hankel example we have: $k = 4$, $f = [\tilde{\alpha} \;\; \tilde{\beta} \;\; \tilde{\gamma} \;\; \tilde{\delta}]^T$, $F_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ and $F_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Formulation (2) then becomes

$$\begin{aligned} &\min_f f^T W f \\ &\text{such that } \;(A + [F_1 f \ldots F_n f])x = b + F_{n+1} f, \end{aligned} \tag{4}$$

with $W \in \mathcal{R}^{k\times k}$ a diagonal weighting matrix.

2.2 STLN

In the STLN formulation the different entries of ΔA are stacked in a vector $\alpha \in \mathcal{R}^{q\times 1}$ while the different entries of Δb (which are not already accounted for in α) are stacked in a vector $\beta \in \mathcal{R}^{d\times 1}$. The structure is memorized in the matrices $X \in \mathcal{R}^{m\times q}$, $F \in \mathcal{R}^{m\times d}$, $G \in \mathcal{R}^{m\times q}$ such that:

$$\begin{aligned} \Delta A x &= X\alpha \\ \Delta b &= F\beta + G\alpha. \end{aligned}$$

For the simple Hankel example this results in: $q = 3$, $d = 1$, $\alpha = [\tilde{\alpha} \;\; \tilde{\beta} \;\; \tilde{\gamma}]^T$, $\beta = \tilde{\delta}$, $F = [0 \;\; 0 \;\; 1]^T$, $G = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $X = \begin{bmatrix} x(1) & 0 & 0 \\ 0 & x(1) & 0 \\ 0 & 0 & x(1) \end{bmatrix}$. The original structured TLS formulation (2) becomes:

$$\begin{aligned} &\min_{\alpha,\beta,x} \frac{1}{2}(\alpha^T D_\alpha^2 \alpha + \beta^T D_\beta^2 \beta) \\ &\text{such that } \;Ax + X\alpha = b + F\beta + G\alpha, \end{aligned} \tag{5}$$

where $D_\alpha \in \mathcal{R}^{q\times q}$ and $D_\beta \in \mathcal{R}^{d\times d}$ represent weighting matrices (e.g. to represent the number of occurrences of the elements of α and β in ΔA and Δb).

2.3 STLS

STLS somewhat differs from the other formulations since it arranges the *different* elements of $S + \Delta S$ (and not ΔS) in a vector $t \in \mathcal{R}^{k\times 1}$. The structure is memorized in fixed basis matrices $T_i \in \mathcal{R}^{m\times(n+1)}$ such that

$$S = T_0 + \sum_{i=1}^{k} T_i s(i)$$

$$S + \Delta S = T_0 + \sum_{i=1}^{k} T_i t(i),$$

with $s, t \in \mathcal{R}^{k\times 1}$. For our small Hankel example this gives: $k = 4$, $s = [5 \quad 6 \quad 7 \quad 8]^T$, $t = [5+\tilde{\alpha} \ \ 6+\tilde{\beta} \ \ 7+\tilde{\gamma} \ \ 8+\tilde{\delta}]^T$, $T_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^T$, $T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^T$, $T_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}^T$, $T_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^T$, $T_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T$. The structured TLS formulation (2) then becomes:

$$\min_{t,y} \sum_{i=1}^{k} W(i,i)(t(i) - s(i))^2$$
$$\text{such that } \ Ty = 0 \ \text{ and } \ y^T y = 1, \tag{6}$$

with $t \in \mathcal{R}^{k\times 1}$ the different elements of $S + \Delta S$, $T = T_0 + \sum_{i=1}^{k} t(i) T_i$ and $W \in \mathcal{R}^{k\times k}$ a diagonal weighting matrix.

2.4 Comparison of Formulations

The multitude of ways for formulating the problem leads to the pertinent question on their usefulness. As we will show now and in the remainder of the paper, the main differences between these formulations are the solution methods you can apply to them. In this view, it should come as no surprise that some formulations lead to more efficient implementations, while others improve our insight in the process of how the solution is derived and what the links are with the unstructured case.

For the time being we indicate the following differences. STLN is for now, the only formulation in which an extension towards multiple right-hand sides has been formulated and implemented [14] successfully. It is also the only formulation that allows in a fairly simple way to minimize the error in another norm than L_2 (e.g. L_1, L_∞ see [11]). Clearly, the extendability of the STLN approach is due to its straightforward approach. This is also the reason why STLN can be extended to deal with nonlinearly structured matrices [12]. If we would try to extend the STLS formulation to nonlinearly related structures, we would need matrix functions T_i, instead of constant matrices T_i, thereby making a derivation similar to that in [4] impossible.

A difference that was already mentioned before, is that STLS parametrizes the corrected matrix $S + \Delta S$ instead of the error matrix ΔS.

3 Algorithms

In this section we describe some algorithms for solving the three problem formulations. Keep in mind however that this is a major field of development and that all of these algorithms still have to prove their usefulness on real-life examples. On simulation examples most of them perform well, provided the matrix sizes are not too large (see next section).

3.1 CTLS

In [1] and [2] it is proven that (4) is equivalent to solving the following minimization problem in x:

$$\min_x \left[\begin{array}{c} x \\ -1 \end{array}\right]^T S^T (H_x W^{-1} H_x^T)^{-1} S \left[\begin{array}{c} x \\ -1 \end{array}\right], \tag{7}$$

with $H_x = \sum_{i=1}^n x(i) F_i - F_{n+1}$. To solve (7) [2] proposes a Newton method with analytically derived gradients and Hessians. However, the multiple inversions and multiplications in the calculation of these gradients and Hessians cause an important loss of accuracy. We will therefore use a quasi-Newton method with BFGS updating rule [7], denoted here as CTLS1.

3.2 STLN

From (5) we see that we have to deal with a constrained optimization problem with a quadratic objective function and nonlinear constraints (the nonlinearity comes from the term $X\alpha$).

In [11] and [14] a weighting method is used to turn the constrained optimization problem into an unconstrained one and afterwards the Gauss-Newton method is applied to this unconstrained nonlinear least squares problem. However, at the starting point the residual of this nonlinear least squares problem is generally not close to zero, which implies (see [7]) that in most practical cases the convergence rate of this algorithm is mostly linear. However, since the use of weighting methods to transform unconstrained problems into constrained ones leads to ill conditioning and does not allow fully automated algorithms (one has to tune the weight every time), we solve the problem as an equality-constrained LS problem [3, Section 5.1], derived from the original formulation by linearizing the nonlinear constraints around the current working point. This method is referred to as STLN1. However, one of the disadvantages of this method is the fact that a lot of iterations are needed before a point is reached where the nonlinear equality constraints from (5) are more or less satisfied (i.e. so that we can say that $S + \Delta S$ is a rank-deficient matrix). As described in [10], one of the main reasons of this shortcoming is the fact that STLN1 uses no curvature information of the constraints. In [10] we propose an improved version of STLN1, denoted STLN2, by taking these curvature terms into account. The price we have to pay for the increased convergence rate (superlinear convergence) is that we use more memory (due to the introduction of second-order information of the constraints), which can dramatically slow down computations.

3.3 STLS

It is proven in [4] that (6) is equivalent to the following formulation: find the solutions u and v according to the smallest τ, such that

$$Sv = D_v u \tau \quad u^T D_v u = 1 \tag{8}$$

$$S^T u = D_u v \tau \quad v^T D_u v = 1, \tag{9}$$

with $D_v = \sum_{i=1}^k \frac{1}{W(i,i)} (T_i v)(T_i v)^T \in \mathcal{R}^{m \times m}$, $D_u = \sum_{i=1}^k \frac{1}{W(i,i)} (T_i^T u)(T_i^T u)^T \in \mathcal{R}^{n \times n}$ and $W \in \mathcal{R}^{k \times k}$ a diagonal weighting matrix. The interesting thing about this formulation is its similarity to the classical Singular Value Decomposition (SVD). Observe that (8)-(9) represents a kind of nonlinear SVD (the nonlinearity is represented by the square matrices

D_u and D_v), called Riemannian SVD (see [5] and [6] for more details).
In [4] an inverse iteration algorithm based on equations (8)-(9) is proposed. We will denote this algorithm by STLS1.

4 Interpretation

The different formulations and their corresponding solution methods have an interesting interpretation which allows to draw a parallel between the TLS method and the structured TLS approach.
If we start from the TLS formulation, we see that this can also be interpreted as a constrained optimization problem and therefore it is also possible to solve the TLS problem by using some optimization routine. That is exactly what happens in STLN1 and STLN2 but then for the structured TLS case.
However, for reasons of computational efficiency, the TLS problem should always be solved via the SVD of the extended data matrix $[A\ \ b]$. That is when STLS enters the picture: as mentioned before, STLS solves the structured TLS by reformulating it as a kind of nonlinear SVD. Nevertheless, STLS is (not yet) the preferred method, because efficient algorithms for solving this nonlinear SVD are not yet developed.
If we examine (7), we see that this is the nonlinear Rayleigh quotient associated to the nonlinear SVD of the STLS formulation (see [9]). As shown in the next section, the squaring effect of S that occurs in the nonlinear Rayleigh quotient, is responsible for the lower accuracy of CTLS (for the same reason one would not use the eigenvalue decomposition (EVD) of S^TS to calculate the singular vectors of S). One should also be careful not to calculate the vector x of (7) by sequentially performing an EVD of $S^T(H_xW^{-1}H_x^T)^{-1}S$ with updating of H_x, since this leads to erroneous results (see [9], [15]).

5 Numerical Example

In this section we test the different described algorithms on the numerical example described in [4]. It is a very small modeling example which reduces to a structured TLS formulation (2) with $A = [6\ \ 5\ \ 4\ \ 3\ \ 2]^T$ and $b = [5\ \ 4\ \ 3\ \ 2\ \ 1]^T$. Observe that the extended data matrix $[A\ \ b]$ has a Hankel structure. The advantage of this small example is that it is possible to calculate the exact solution of the structured TLS problem (see e.g. [14]). As starting value for ΔS we take the zero matrix and for x we take $7.807764064044154e-01$.
In Figure 4 we see $\|x - x_{opt}\|_2$ (where x_{opt} stands for the exact solution of x, $x_{opt} = 7.629230150743218e-01$) as a function of the number of iterations. In this figure we see that CTLS1 indeed suffers from a loss of accuracy, as predicted in the previous section. Furthermore we observe that STLN1 and STLS1 have more or less linear rates of convergence, while STLN2 converges in a superlinear fashion. Table 1 shows the number of flops required by the different algorithms applied on this small modeling example. We say that the algorithm converges if x reaches a steady state.

TABLE 1

Number of flops needed by the different algorithms described in this paper, applied to the small modeling example described in Section 5.

	CTLS1	STLN1	STLN2	STLS1
number of flops	64089	54512	17787	34213

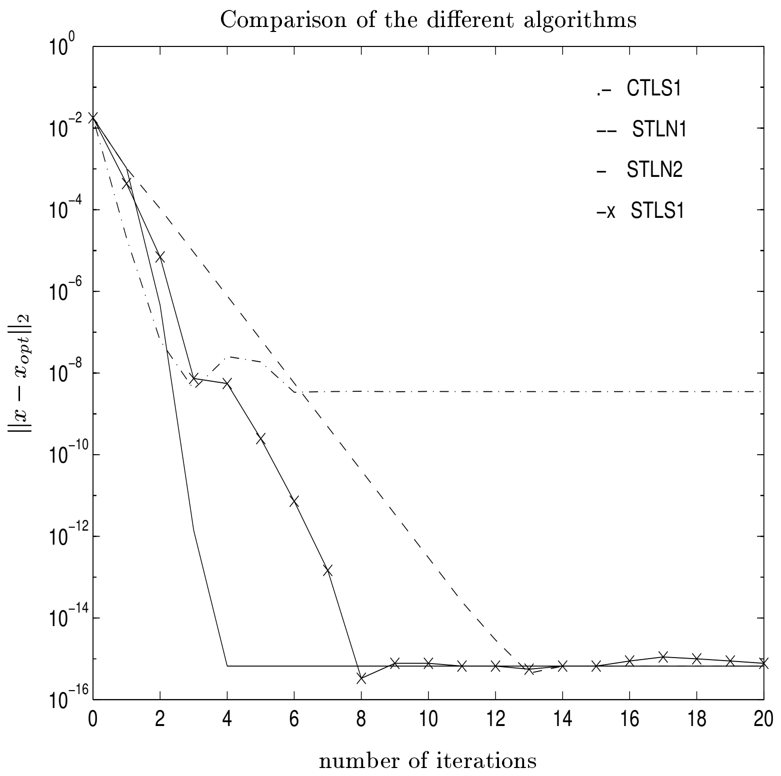

FIG. 4. $\|x - x_{opt}\|_2$ as a function of the iteration number for the described algorithms: CTLS1, STLN1, STLN2 and STLS1.

6 Conclusions and Future Research

In many applications of system identification, signal processing,... the kernel problem is a typical structured TLS problem. Different strategies are presented here to tackle the structured TLS, ranging from pure optimization routines (STLN) to nonlinear EVD (CTLS) or nonlinear SVD (STLS). Right now, the STLN approach is the most viable. However, other methods (CTLS and STLS) highlight the link that exists between TLS and structured TLS.

Future research will be focused on several points. First of all, the optimization methods that are used by STLN need to be fine-tuned for large real-life applications. One very good candidate for this task is the conjugate gradient algorithm (low memory usage) applied to an augmented Lagrangian function, which takes the constraints into account. Secondly, better alternatives to the inverse iteration algorithm for STLS have to be developed. It is expected that the inverse iteration algorithm for nonlinear SVD can be extended towards an inverse iteration algorithm with a shift, even a variable shift (Rayleigh quotient iteration).

References

[1] T. J. Abatzoglou and J. M. Mendel, *Constrained Total Least Squares*, in Proc. IEEE International Conf. on Acoustics, Speech & Signal Processing, Dallas, 1987, pp. 1485–1488.

[2] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, *The Constrained Total Least Squares Technique and its Applications to Harmonic Superresolution*, IEEE Trans. on Signal Processing, SP-39 (1991), pp. 1070–1086.

[3] Åke Björck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1996.

[4] B. De Moor, *Total least squares for affinely structured matrices and the noisy realization problem*, IEEE Trans. on Signal Processing, SP-42 (1994), pp. 3004–3113.

[5] B.L.R. De Moor, *The Riemannian Singular Value Decomposition*, in SVD and Signal Processing III: algorithms, architectures, and applications, M. Moonen and B. De Moor, eds., Elsevier Science B.V., Amsterdam, 1995.

[6] B. De Moor, *Engineering applications for structured total least squares and the Riemannian SVD*, in this volume, S. Van Huffel, ed., SIAM, Philadelphia, PA, 1997.

[7] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, New York, 1987.

[8] R. Fletcher, *Methods for Nonlinear Constraints*, in Nonlinear Optimization, M.J.D. Powell, ed., Nato Conference Series , Cambridge, Great Britain, 1982, pp.185–211 .

[9] P. Lemmerling, B. De Moor, and S. Van Huffel, *On the equivalence of constrained total least squares and structured total least squares*, IEEE Trans. on Signal Processing, SP-44 (1996), no. 11, to appear.

[10] P. Lemmerling and S. Van Huffel, *Superlinear algorithms for solving the structured total least norm problem*, ESAT/SISTA report TR 95-44, Dept. of Electr. Eng., Katholieke Universiteit Leuven, Belgium, 1995.

[11] J.B. Rosen, H. Park, and J. Glick, *Total least norm formulation and solution for structured problems*, SIAM Matrix Anal. Appl., 17 (1996), pp. 110–126.

[12] J.B. Rosen, H. Park, and J. Glick, *Structured total least norm techniques for linearly and nonlinearly structured problems*, in this volume, S. Van Huffel, ed., SIAM, Philadelphia, PA, 1997.

[13] S. Van Huffel and J. Vandewalle, *The total least squares problem : computational aspects and analysis*, SIAM, Philadelphia, PA, 1991.

[14] S. Van Huffel, H. Park, and J. Ben Rosen, *Formulation and Solution of Structured Total Least Norm Problems for Parameter Estimation*, IEEE Trans. on Signal Processing, SP-44 (1996), no. 10, to appear.

[15] S. Van Huffel, B. De Moor, and H. Chen, *Relationships between Structured and Constrained TLS, with applications to signal enhancement.* In Systems and Networks: Mathematical Theory

and Applications, U. Helmke, R. Mennicken, J. Saurer, eds., Vol. II, Akademie Verlag, Berlin (Mathematical Research: Vol. 79), 1994, pp. 899–904.