Relationships between Structured and Constrained TLS, with applications to Signal Enhancement

Sabine Van Huffel, Bart De Moor[1] and Hua Chen
ESAT Lab., Department of Electrical Engineering, Katholieke Universiteit Leuven, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium (e-mail: sabine.vanhuffel@esat.kuleuven.ac.be, bart.demoor@esat.kuleuven.ac.be, hua.chen@esat.kuleuven.ac.be).

1 Introduction: the CTLS and STLS problem.

A problem that often occurs in system identification, modal analysis, signal processing [6], etc. is to approximate a given affinely structured data matrix A (e.g. Hankel, Toeplitz,...) by one of lower rank having the *same* structure. Since there is linear dependence among the error entries in A, the Total Least Squares (TLS) solution no longer yields optimal statistical estimators [5]. To get more accurate estimates, Abatzoglou and Mendel [1] extended the classical TLS problem to incorporate the algebraic dependence of the errors in A and called their extension "constrained TLS".

__Definition:__ **Constrained TLS (CTLS) problem** [1]. *Let $A_{p\times q}$ be given and $\Delta A_j = F_j\delta a$ be the correction on its jth column with F_j a fixed $p\times m$ matrix and δa a zero-mean white noise vector of minimal dimension m. The CTLS problem seeks to ("$*$" denotes the conjugated transpose):*

$$\min_{\delta a, y} \|\delta a\|_2^2 \text{ subject to } \{A + [F_1\delta a, \ldots, F_q\delta a]\}y = 0 \text{ and } y^*y = 1 \tag{1}$$

e.g. if A is Hankel then $F_j = [0_{p\times(j-1)}\ I_p\ 0_{p\times(q-j)}]$ with $m = p+q-1$. For simplicity, we assume that δa is white (if not, a whitening transformation can always be applied [1]). As shown in [1], the CTLS solution $\widehat{y}$ can be obtained from the following unconstrained minimization problem:

$$\min_{y, y^*y=1} y^*A^*D_y^{-1}Ay \text{ with } D_y = H_yH_y^* \text{ and } H_y = \sum_{j=1}^{q} y_jF_j \tag{2}$$

if H_y has full rank and $p \leq m$. The corresponding correction matrix $\Delta A = [F_1\delta a, \ldots, F_q\delta a]$ with $\delta a = -H_{\widehat{y}}^{\dagger}A\widehat{y}$("$\dagger$" denotes the pseudo-inverse). In [1], a complex version of Newton's method has been derived and applied to find the CTLS solution. However, convergence problems occur as soon as the noise in A is no longer small which is the case in biomedical signal processing [6]. Better convergence results have been obtained with substantially less computations by solving (2) simply by means of inverse iteration, i.e. for $k = 1, 2, \ldots$ up to convergence, solve for $y(k)$:

$$R^*Q_1^*D_{y(k-1)}^{-1}Q_1Ry(k) = y(k-1)$$

[1]The first two authors are Research Associates of the Belgian NFWO (National Fund for Scientific Research); support provided by the Belgian Program on Interuniversity Attraction Poles (IUAP-50).

with $A = Q_1 R$, $Q_1^* Q_1 = I_q$ and R upper triangular, the truncated QR decomposition, and $y(0)$ an initial estimate of the eigenvector corresponding to the minimal eigenvalue of $A^* D_{\tilde{y}}^{-1} A$. Typically, we take as $y(0)$ the eigenvector corresponding to the minimal eigenvalue of $A^* A$. Convergence is obtained as soon as $\|y(k) - y(k-1)\|_2$ is below a given tolerance ϵ. This CTLS problem is a special case of the following:

Definition: Structured TLS (STLS) problem [3]. *Given an $m \times 1$ data vector a and weight vector w. Let $B(b) = B_0 + b_1 B_1 + \ldots + b_n B_n$ be an affine matrix function of the parameter vector b where $B_j, j = 0, 1, \ldots, n$ are fixed given $p \times q$ matrices. Find a rank-deficient matrix in the affine set $B(b)$ such that a given quadratic function $[a, b, w]_2^2$ of the parameters b_i is minimized, i.e.*

$$\min_{b,y} [a, b, w]_2^2 \text{ subject to } B(b)y = 0 \text{ and } y^* y = 1 \tag{3}$$

e.g. if we want to approximate a $p \times q$ Hankel matrix A by one of lower rank, say $B(b) = \sum_{i=1}^n b_i B_i$, then $a = [A_{11}, \ldots, A_{p1}, A_{p2}, \ldots, A_{pq}]$ contains the $n = p + q - 1$ different entries of $A = \sum_{i=1}^n a_i B_i$, $w = [1, \ldots, 1]^T$ if all entries are equally weighted. $B_0 = 0$ while B_j is a zero matrix with ones on its jth antidiagonal containing the entries $(i, j-i+1), i = 1, \ldots, j$. Here, we focus on the STLS problem (3) with $m = n$:

$$\min_{b,y} \|a - b\|_2^2 \text{ subject to } B(b)y = 0 \text{ and } y^* y = 1 \tag{4}$$

in which weights w are not considered for simplicity.

As proven in [4, 3, Theorem 1] by making use of Lagrange multipliers, the STLS solution is obtained from the following nonlinear generalized Singular Value Decomposition (SVD) problem:
Find the triplet (u, τ, v) corresponding to the minimal τ that satisfies

$$Av = D_v u \tau \quad u^* D_v u = 1 \quad \text{and} \quad A^* u = D_u v \tau \quad v^* D_u v = 1 \tag{5}$$

D_u, defined by $\sum_{i=1}^m B_i^* (u^* B_i v) u = D_u v$, and D_v, defined similarly by $\sum_{i=1}^m B_i (u^* B_i v) v = D_v u$, are Hermitian nonnegative definite with elements that are quadratic in the components of u and v. The STLS solution y is given by $y = v / \|v\|_2$ and the components of b are obtained from $b_k = a_k - u^* B_k v \tau$, $k = 1, \ldots, m$.

The following properties of the STLS solution are important for our comparison:
1. If $B_0 = 0$ and $W = \mathrm{diag}(w_i)$ then $(a - b)^* W b = 0$, i.e. the residual vector $a - b$ is orthogonal to b. This property is necessary but not sufficient for optimality.
2. If D_v in (5) is invertible then the STLS solution also follows from the following equivalent minimization problem [4]:

$$\min_v \tau^2 = \min_v v^* A^* D_v^{-1} A v \text{ subject to } v^* D_u v = 1 \tag{6}$$

3. The object function (6) is scaling invariant, i.e. (6) does not change if we replace v by v/α.
4. With D_u and D_v as defined in (5): $u^* D_v u = v^* D_u v$ for arbitrary u and v.
5. If (u, τ, v) satisfies (5) then any triplet $(\frac{u}{\alpha}, \tau, \frac{v}{\beta})$ with $\alpha\beta = 1$ also satisfies (5), i.e. the vectors u and v are not unique.

A straightforward linear convergent algorithm that solves the set of nonlinear equations (5) is outlined in [3, 4] and is based on the inverse iteration method to find the smallest singular value and corresponding singular vectors of a matrix. For given D_u and D_v, one step of inverse iteration is performed, using the QR decomposition of A, in order to obtain new estimates of u and v. These are then used in updating D_u and D_v, etc. Convergence is achieved when $\|B(b)v\|_2$ and $\|B(b)^*u\|_2$ (with $B(b)$ the current reconstructed approximation matrix) fall below a given tolerance ϵ.

Many STLS applications are treated in [3]. Here, we confine our attention to a comparison of STLS with CTLS and other suboptimal approaches described in [6, 2]. Section 2 then compares the different methods in a particular signal enhancement problem in which we try to approximate a Hankel matrix by one of lower rank.

The **suboptimal methods** [2, 6] try to solve the minimization problem:

$$\min_{B \in S_r} \|A - B\|_F \tag{7}$$

where S_r is the set of rank r matrices having the same structure as A, by decomposing this problem in 2 simpler subproblems. First, one gets the rank r approximation $A_{(r)}$ of A. Hereto, Cadzow [2] arranges the data a in a matrix $A_{p\times q}$ with p, q determined by the considered application (e.g. $p = q(+1)$ in Sec.2) and computes the rank r truncated SVD of A, i.e. given the SVD $A_{p\times q} = \sum_{i=1}^{q} \sigma_i u_i v_i^*$ with $\sigma_i \geq \sigma_{i+1}, \forall i$, then $A_{(r)} = \sum_{i=1}^{r} \sigma_i u_i v_i^*$. As shown in [6], better resolution properties using noisy data are obtained by first arranging the data in a matrix $A_{p\times q}$, where $p >> q$, and then computing the Minimum Variance (MV) estimate $A_{(r)} = \sum_{i=1}^{r} (\sigma_i^2 - p\sigma_\nu^2)\sigma_i^{-1} u_i v_i^*$ of the noise-free data matrix. σ_ν^2 is the (estimated) variance of the complex noise added to the given data a. Since $A_{(r)}$ destroys the structure of A we next recover the required structure and this completes the first iteration, e.g. if A is Hankel (as in Sec.2), the closest Hankel matrix is simply obtained by replacing the antidiagonals of $A_{(r)}$ by the average of their elements. However, the new structured matrix is no longer of rank r. One then iterates by again computing the truncated SVD [2] or MV estimate [6] and restoring the required structure, etc. This process converges but not (necessarily) to the minimum of (7). In fact, it is shown in [3, 4] that these suboptimal approaches do not deliver an L_2-optimal solution since the orthogonality property is not satisfied, i.e. $\|(A - B)^* B\|_F = (a - b)^* W b \neq 0$ where W is a weighting matrix (determined by the structure of A, e.g. if A is Hankel then $W = \mathrm{diag}(1, \ldots, q-1, q, \ldots, q, q-1, \ldots, 1)$).

Let's now **compare STLS with CTLS**. CTLS, as well as STLS, allow to approximate given structured matrices $A = B_0 + \sum_{i=1}^{n} a_i B_i$, such as Toeplitz and Hankel matrices or matrices with certain zero patterns or error-free entries, by rank-deficient matrices $B(b)$ having the same structure. In these cases, A and B have the *same* structure implying that $m = n$ and both formulations coincide if $[a, b, w]_2^2 = \|a - b\|_2^2$. Although both formulations allow to introduce weighting matrices $W_{n\times n}$, we focus here on the unweighted case, i.e. $W = I_n$. Define $\delta a = -(a - b)$ the correction vector then (4) can always be written as (1) where $\Delta A = [F_1 \delta a, \ldots, F_q \delta a] = \sum_{i=1}^{n} \delta a_i B_i = -\sum_{i=1}^{n} a_i B_i + \sum_{i=1}^{n} b_i B_i = -(A - B(b))$, i.e. the jth column $F_j \delta a$ of ΔA equals the jth column of $\sum_{i=1}^{n} \delta a_i B_i$ and this defines the different F_j as used in the CTLS formulation. Differences between both formulations arise when A and $B(b)$ or $[F_1\delta a, \ldots, F_q\delta a]$ are structured in a *different* way, e.g. when A is an arbitrary matrix and B is the closest rank-deficient Hankel matrix. In that case, $A - B$ is not a Hankel matrix and we can not find F_j and δa such that CTLS formulates exactly the same problem. In addition, the STLS problem formulation is more general in the sense that more general quadratic criteria than the 2-norm can be used and other constraints can be easily added merely because the solution of the STLS problem is based on the use of Lagrange multipliers, e.g. if one wants to find a matrix B closest to A such that B has a specified singular value β, the STLS problem formulation is easily found [3]:

$$\min_{b,u,v} \|a - b\|_2^2 \text{ subject to } Bv = u\beta,\ B^*u = v\beta \text{ and } u^*u = 1$$

Comparing the way in which CTLS and STLS solve their problem, the following conclusions can be made. First, $D_y = H_y H_y^*$, used in the CTLS problem (2), equals D_v, used in the STLS problem (6), for $y = v$, e.g. if $A_{p\times q}$ is Hankel then $H_y = \begin{bmatrix} y_1 & \cdots & y_q & & 0 \\ & \ddots & & \ddots & \\ 0 & & y_1 & \cdots & y_q \end{bmatrix}_{p\times(p+q-1)} = \sum_{i=1}^{q} y_i F_i$. Second, $D_u = H_u H_u^*$, used to solve the STLS problem (5) and (6), is a $q \times q$ Hermitian positive definite Toeplitz matrix generated in the same way from the elements of u as $D_y = H_y H_y^*$ from the elements of y. Observe that D_u is not used in the CTLS problem which is the main reason why the CTLS algorithm, outlined above, is simpler than the STLS algorithm based on inverse iteration. Comparing (2) with (6), we observe that the main difference between both solutions is the fact that the STLS scaling constraint $v^* D_u v = 1$ has been replaced by a simpler constraint $v^*v = 1$ in the CTLS problem (as formulated here) or the requirement that one of its components be ± 1, e.g. $v_q = -1$ as used in [1]. This is the main reason why the CTLS algorithms, presented here and in [1], as well as the classical algorithms Steiglitz-McBride and Iterative Quadratic Maximum Likelihood, do not converge to the L_2-optimal solution, as shown with a simple counterexample in [4] (despite misleading claims in the literature). Indeed, the scaling or normalization of v in (6) has no effect on $v^* A^* D_v^{-1} A v$ but one may not replace the constraint $v^* D_u v = 1$ by another one, e.g. $v^*v = 1$, as done in (1), or $v_q = -1$ [1]. The appearance of D_u is a crucial requirement for the optimal solution.

2 Results

Here, we compare the performance of the STLS and CTLS algorithms with that of the suboptimal approaches [2, 6] in the following exponential data modeling problem. Given N complex data points $x_n, n = 0, \ldots, N-1$, modeled as:

$$x_n = \sum_{k=1}^{K} c_k z_k^{t_n} = \sum_{k=1}^{K} (a_k e^{j\phi_k}) e^{(-d_k + j2\pi f_k)t_n} \quad n = 0, \ldots, N-1 \tag{8}$$

where $j = \sqrt{-1}$ and t_n is the time lapse between the effective time origin and sample x_n, the objective is to estimate the frequencies f_k, damping factors d_k, amplitudes a_k

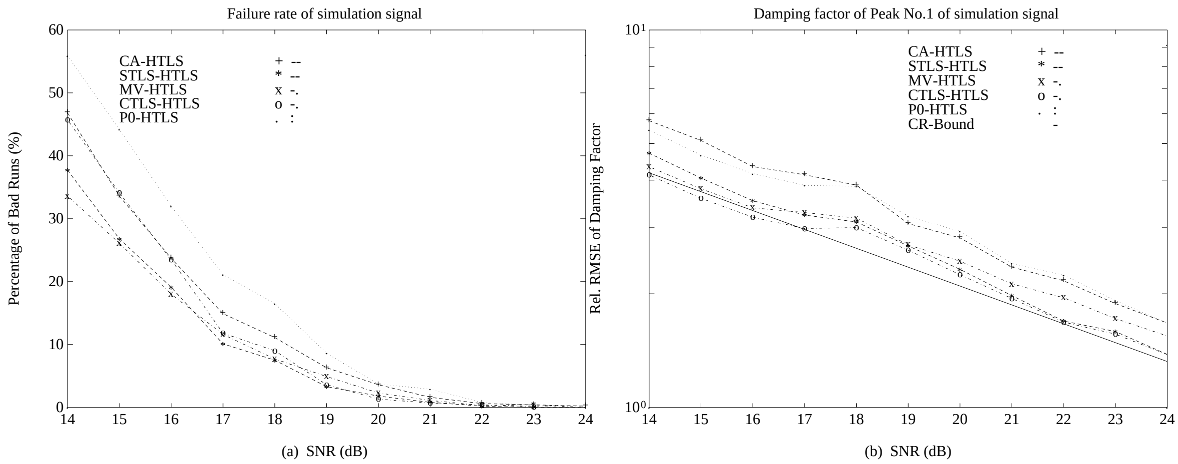

Figure 1: The results (averaged over 1000 runs) of applying STLS-HTLS, CTLS-HTLS, CA-HTLS, MV-HTLS and the nonenhanced P0-HTLS to the simulation signal $x_n = e^{(-0.01+j2\pi 0.2)n} + e^{(-0.02+j2\pi 0.22)n}$, $n = 0, \ldots, 24$ ($K = 2$, $N = 25$) [6] versus the peak signal-to-noise ratio $= 10 \log(1/\sigma_\nu^2)$ where σ_ν^2 is the variance of the added complex Gaussian noise: (a) % of times that each method fails to resolve the 2 peaks within the frequency intervals 0.2 ± 0.0094Hz and 0.22 ± 0.0106Hz; (b) Relative root mean-squared errors (excluding failures) obtained for estimating $d_1 = 0.01$Hz.

and phases ϕ_k, $k = 1, \ldots, K$. Prior to parameter estimation, the data are enhanced as follows. Arrange the x_n in a $p \times q$ Hankel matrix A in which $N = p + q - 1$. Except for Cadzow's method in which $p = q$ is taken, q is set to $K + 1$. Approximate now A by the closest rank K Hankel matrix B by means of the CTLS algorithm based on inverse iteration, the STLS algorithm of [3], the MV method [6] and Cadzow's method [2]. The cleaned-up data samples $\widehat{x}_n$, $n = 0, \ldots, N-1$, which are found along the first column and last row of the resulting B, are then used to estimate the required signal parameters f_k, d_k, a_k, ϕ_k in (8) by an improved version of Kung's state-space method based on TLS and called here HTLS. See [6] for more details. If respectively no signal enhancement, Cadzow's method or the MV method with one iteration, CTLS or STLS with $\epsilon = 10^{-7}$, is used prior to HTLS, the method is called P0-HTLS, CA-HTLS, MV-HTLS, CTLS-HTLS or STLS-HTLS.

Simulations show that MV-HTLS significantly improves the *resolution* of interfering peaks. As shown in Fig. 1(a), the resolution is doubled compared to Cadzow's method that still performs better than the nonenhanced method P0-HTLS. CTLS-HTLS and STLS-HTLS do not exhibit good resolution properties except on small sampled problems, e.g. $N = 25$ as shown in Fig.1. For larger N, examples are found in which the resolution of STLS-HTLS and CTLS-HTLS is even worse than for the nonenhanced method P0-HTLS, e.g when solving [6, Example 1] where $N = 128$ and $K = 5$. However, if the peaks are resolved, then the obtained STLS-HTLS and CTLS-HTLS parameter estimates f_k, d_k, a_k, ϕ_k are clearly more accurate (closer to the Cramér-Rao bound) than the other estimates, as illustrated in Fig.1(b). On the other hand, STLS-HTLS and CTLS-HTLS, requiring typically 50 to 100 iterations or more, are computationally much more expensive than the suboptimal approaches that only

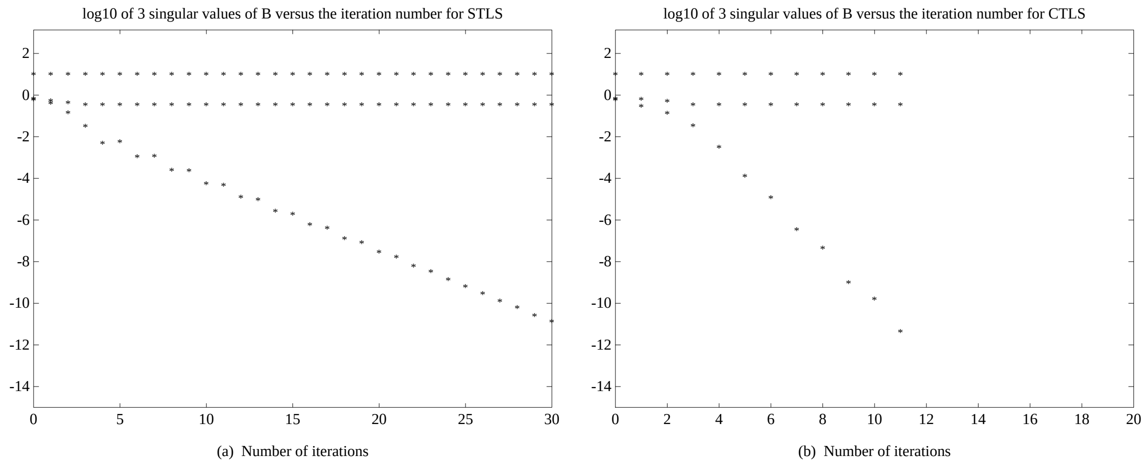

Figure 2: Linear convergence behavior of the (a) STLS versus (b) CTLS algorithm for one run on the simulation signal used in Fig.1. $\sigma_\nu = 0.01$. $\epsilon = 10^{-10}$. The logarithm of the 3 singular values of the current approximation matrix $B(b)$ are plotted versus the number of performed iterations.

perform one iteration here. Besides a considerably smaller number of computations per iteration step, the CTLS algorithm converges clearly faster, usually a factor 2, than the STLS algorithm (see Fig.2). Convergence problems, especially with STLS, occur for large noise levels and large N (compared to K). The convergence properties of these algorithms are not yet fully understood and are currently under study.

References

[1] T.J. Abatzoglou, J.M. Mendel and G.A. Harada, *The constrained total least squares technique and its application to harmonic superresolution.* IEEE Trans. on SP 39, 1991, pp.1070-1086.

[2] J. Cadzow, *Signal enhancement: a composite property mapping algorithm.* IEEE Trans. on ASSP 36, 1988, pp. 49-62.

[3] De Moor B. *Structured total least squares and L_2 approximation problems.* Lin. Alg. Appl. 188, July 1993, to appear.

[4] De Moor B. *Total least squares for affinely structured matrices and the noisy realization problem.* ESAT-SISTA report TR-93-21, ESAT Lab., K.U.Leuven, Belgium, May 1993.

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

[6] S. Van Huffel, *Enhanced Resolution Based on Minimum Variance Estimation and Exponential Data Modeling.* Signal Processing 33 (3), August 1993, to appear.