95_08

Formulation and Solution of Structured Total Least Norm Problems for Parameter Estimation

Sabine Van Huffel, *Member, IEEE*, Haesun Park, *Member, IEEE*, and J. Ben Rosen

***Abstract*—The total least squares (TLS) method is a generalization of the least squares (LS) method for solving overdetermined sets of linear equations $Ax \approx b$. The TLS method minimizes $\|[E|-r]\|_F$, where $r = b - (A + E)x$, so that $(b - r) \in Range(A + E)$, given $A \in \mathbf{C}^{m\times n}$, with $m \geq n$ and $b \in \mathbf{C}^{m\times 1}$. The most common TLS algorithm is based on the singular value decomposition (SVD) of $[A|b]$. However, the SVD-based methods may not be appropriate when the matrix A has a special structure since they do not preserve the structure. Recently, a new problem formulation known as structured total least norm (STLN), and the algorithm for computing the STLN solution, have been developed. The STLN method preserves the special structure of A or $[A|b]$ and can minimize the error in the discrete L_p norm, where $p = 1$, 2 or ∞. In this paper, the STLN problem formulation is generalized for computing the solution of STLN problems with multiple right-hand sides $AX \approx B$. It is shown that these problems can be converted to ordinary STLN problems with one right-hand side. In addition, the method is shown to converge to the optimal solution in certain model reduction problems. Furthermore, the application of the STLN method to various parameter estimation problems is studied in which the computed correction matrix applied to A or $[A|B]$ keeps the same Toeplitz structure as the data matrix A or $[A|B]$, respectively. In particular, the L_2 norm STLN method is compared with the LS and TLS methods in deconvolution, transfer function modeling, and linear prediction problems and shown to improve the accuracy of the parameter estimates by a factor of 2 to 40 at any SNR.**

I. Introduction
Structured Total Least Norm (STLN) Problems

THE total least squares (TLS) method is a frequently adopted data fitting technique in signal processing applications for solving an overdetermined set of linear equations $Ax \approx b$ [12], [13], [19], [27]. The TLS problem can be stated as that of finding E and x such that

$$\min_{E,x} \|[E|-r]\|_F \tag{1}$$

where $r = b - (A + E)x$, given $A \in \mathbf{C}^{m\times n}$, with $m \geq n$, and $b \in \mathbf{C}^{m\times 1}$. This problem allows the possibility of error in the elements of a given matrix A so that the modified matrix is given by $A + E$, where E is a correction matrix to be determined. The TLS solution is commonly obtained from the singular value decomposition (SVD) of $[A|b]$ [27] and is able to compute a strongly consistent estimate of the true solution x when the errors in the data matrix $[A|b]$ are rowwise independently and identically distributed with zero mean and equal variance. If, in addition, the errors are normally distributed, the TLS solution has maximum likelihood properties. However, these nice properties are lost in applications where the matrix A has a special structure since then, the errors are (non)linearly related, and the TLS method does not preserve this special structure. In fact, the matrix E obtained from the SVD will typically be dense, with no structure, even when A is structured or sparse.

Recently, a new problem known as structured total least norm (STLN) has been presented and can be formulated as follows [20] and [21]. Consider a structured matrix $A \in \mathbf{C}^{m\times n}$, of which $q(\leq mn)$ different elements are subject to error, and define a vector $\alpha \in \mathbf{C}^q$ representing the corresponding elements of the structured correction matrix E. If many elements of E must have the same value, then q is the number of *different* such elements. The vector α and the matrix E are equivalent in the sense that given E, α is known, and vice versa. Then, the residual vector $r = b - (A + E)x$, is a function of (α, x). Let $D \in \mathbf{R}^{q\times q}$ be a diagonal weighting matrix that accounts for the repetition of elements of α in the matrix E. Then, the STLN problem can be stated as

$$\min_{\alpha,x} \left\| \begin{bmatrix} r(\alpha, x) \\ D\alpha \end{bmatrix} \right\|_p \tag{2}$$

where $\|\cdot\|_p$ is the vector p norm for $p = 1$, 2, or ∞. For $p = 2$ and a suitable choice for D, problem (2) is equivalent to the TLS problem (1), with the additional requirement that the matrix E must have the given structure.

Algorithms for computing the STLN solution have been developed [20], [21]. The STLN method preserves any given affine structure in E or $[E|-r]$ and can minimize these corrections in the discrete L_p norm, where $p = 1, 2$ or ∞ [21]. A theoretical justification and computational testing of

Manuscript received February 23, 1995; revised April 16, 1996. This paper is supported by the Belgian Programme on Interuniversity Poles of Attraction (IUAP-nr.17), initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, by the European Community 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." H. Park is supported, in part, by the National Science Foundation grants CCR-9209726 and CCR-9509085. J. B. Rosen is supported, in part, by the National Science Foundation grant CCR-9509085, an ARPA/NIST grant, and the Minnesota Supercomputer Institute. The associate editor coordinating the review of this paper and approving it for publlication was Prof. Michael D. Zoltowski.

S. Van Huffel is with the Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, K. Mercierlaan 94, 3001 Leuven-Heverlee, Belgium.

H. Park is with the Computer Science Department, University of Minnesota, Minneapolis, MN 55455 USA.

J. B. Rosen is with the Computer Science Department, University of Minnesota, Minneapolis, MN 55455 and the Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA 92093 USA.

Publisher Item Identifier S 1053-587X(96)07134-6.

1053–587X/96$05.00 © 1996 IEEE

the STLN algorithm confirm that it is an efficient method for solving structured problems or when errors can occur only in some of the elements of A or $[A|b]$ [21]. For instance, when the matrix A has Toeplitz or Hankel structure, the new approach guarantees that E has the same Toeplitz structure as A or that only those elements of E that correct possible noisy elements of A are permitted to be nonzero. For earlier work on problems in which the errors occur in a restricted manner, see [5] and [26], and for the use of L_1 norm in total approximation, see [17] and [30]. The extension of the TLS solution to incorporate the algebraic pattern of the errors in A is also studied in [1] and [6] but both for the L_2 norm only. For a comparison of these two algorithms, see [28]. The approach in our algorithm is different from these two, as will become clear later on.

Our study of STLN problems (in several norms) is motivated by the numerical and theoretical results reported in many papers (see e.g., [25] and [27] for an overview), which show that a total least norm fitting ($p = 1, 2, \infty$) does give better results than ordinary least norm fitting (in which only b is corrected) when all data in $[A|b]$ have errors. A strategy for choosing p is not simple, and some discussion on this topic is included in [25]. In particular, for $p = 1$, it is well known that the solution of the ordinary least 1-norm problem $\min \|Ax - b\|_1$ is less sensitive to outliers than the solution with other norms. A theoretical analysis in [24] shows that the robustness of least norm fitting and total least norm fitting is similar for $p = 1$, implying that the TLN solution for $p = 1$ is also less sensitive to outliers than the TLN solution for other norms (e.g., $p = 2, \infty$). This robustness of the TLN problem for $p = 1$ compared with that of the TLN problem for $p = 2$ is also shown for structured problems (i.e., for the STLN problem considered here) in our preliminary experiments, as reported in a recent paper [18]. This is one important aspect of the STLN formulation, in addition to the fact that it can preserve the problem structure.

This paper extends the study of the STLN problem presented in [20] and [21] in the following way. First of all, we show, in Section IV, how to generalize the STLN problem formulation to structured TLN problems with multiple right-hand side vectors $AX \approx B$ and how to compute the corresponding TLN solution. This is an important feature that allows us to solve structure-preserving rank reduction problems for any given rank deficiency. On the other hand, the methods presented in [1] and [6] cannot be directly used for problems with multiple right-hand sides. For clarity of exposition, we briefly describe the STLN algorithms for structured A and $[A|b]$, as presented in [20] and [21], in Sections II and III. Second, in Section V, convergence properties in model reduction problems are discussed, and finally, in Section VI, the STLN method is applied to various problems in system identification and signal processing in which the data matrix is Toeplitz, and the correction matrix E or $[E|-r]$ preserves the same structure. The presented results also hold for Hankel structures since Hankel matrices simply transform to Toeplitz matrices by permutations. Although STLN has this flexibility of allowing any vector norm, we focus in this paper on its two-norm solutions, mainly because this norm is most often used in signal processing. In addition, this norm gives a fair comparison to the currently used LS and TLS methods in the applications under study.

II. STLN Algorithm for Structured Matrix A

Construction of the Matrix X_E: In the iterative algorithm for solving the STLN problem [21], the vector Ex is represented in terms of α. This is accomplished by defining a matrix $X_E \in \mathbf{C}^{m \times q}$ such that $X_E \alpha = Ex$. The matrix X_E consists of the elements of x, with suitable repetition, giving X_E a special structure. In fact, E and X_E have exactly the same number of nonzero elements. The construction of X_E (starting with a zero matrix) is carried out according to the following rule:

> If α_k is the (i, j)th element of E, then x_j is the (i, k)th element of X_E, where $i = 1, \cdots, m$, $j = 1, \cdots, n$, and $k = 1, \cdots, q$.

For example, assume that all six diagonals of a 4×3 Toeplitz matrix are subject to error, and the matrix E is also Toeplitz with first column $[\alpha_3 \quad \alpha_4 \quad \alpha_5 \quad \alpha_6]^T$ and first row $[\alpha_3 \quad \alpha_2 \quad \alpha_1]$. If we define α as $\alpha = [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5 \quad \alpha_6]^T$, then

$$X_E = \begin{pmatrix} x_3 & x_2 & x_1 & 0 & 0 & 0 \\ 0 & x_3 & x_2 & x_1 & 0 & 0 \\ 0 & 0 & x_3 & x_2 & x_1 & 0 \\ 0 & 0 & 0 & x_3 & x_2 & x_1 \end{pmatrix}. \tag{3}$$

Note that when we choose the order of the components of the vector α as $\alpha = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n \quad \alpha_{n+1} \quad \cdots \quad \alpha_{m+n-1}]^T = [E(1, n) \quad E(1, n-1) \quad \cdots \quad E(1, 1) \quad E(2, 1) \quad \cdots \quad E(m, 1)]^T$, then X_E is Toeplitz whenever E is Toeplitz.

STLN Algorithm: In solving the minimization problem (2), a linear approximation to $r(\alpha, x)$ is used [21]. Let Δx represent a small change in x, and let ΔE represent a small change in the variable elements of E. Then, we have $X_E \Delta\alpha = (\Delta E)x$, where $\Delta\alpha$ represents the corresponding small change in the elements of α. Ignoring the second-order terms in $\|\Delta\alpha\|$ and $\|\Delta x\|$, we obtain

$$\begin{aligned} r(\alpha + \Delta\alpha, x + \Delta x) &= b - (A + E + \Delta E)(x + \Delta x) \\ &\approx r(\alpha, x) - X_E \Delta\alpha - (A + E)\Delta x. \end{aligned} \tag{4}$$

The linearization of (2) now becomes

$$\begin{aligned} &\min_{\Delta\alpha, \Delta x} \left\| M \begin{pmatrix} \Delta\alpha \\ \Delta x \end{pmatrix} + \begin{pmatrix} -r \\ D\alpha \end{pmatrix} \right\|_p, \\ &M = \begin{pmatrix} X_E & A + E \\ D & 0 \end{pmatrix}. \end{aligned} \tag{5}$$

To start the iterative algorithm, the initial values of $E = 0$, and the least norm value of $x = x_{ln}$ can be used, where x_{ln} is given by

$$\min_x \|b - Ax\|_p. \tag{6}$$

The STLN algorithm is summarized in Algorithm STLN, which follows:

Algorithm STLN:

Input—A structured total least norm problem (2), with specified matrices A, D, vector b, and tolerance ϵ
Output—Correction matrix $[E|-r]$ and STLN solution x
Begin
1) Set $E = 0$, $\alpha = 0$, compute x from (6) and X_E from x, and set $r = b - Ax$.
2) **repeat**
a) $\min_{\Delta x, \Delta\alpha} \left\| \begin{pmatrix} X_E & A+E \\ D & 0 \end{pmatrix} \begin{pmatrix} \Delta\alpha \\ \Delta x \end{pmatrix} + \begin{pmatrix} -r \\ D\alpha \end{pmatrix} \right\|_p$.
b) Set $x := x + \Delta x$, $\alpha := \alpha + \Delta\alpha$.
c) Construct X_E from x and E from α. Compute $r = b - (A+E)x$.
until $(\|\Delta x\|, \|\Delta\alpha\| \le \epsilon)$
End

For $p = 2$, the LS problem (5) can be solved efficiently by a QR factorization of the matrix M when $A + E$ has full column rank since M has full column rank in this case. For a general matrix M, the standard QR decomposition algorithms require $O\left[(m+q)(n+q)^2\right]$ flops. When A is Toeplitz, a fast triangularization of M can be performed in $O(mn^2 + m^2)$ flops, as described in [20, Section V]. Note that when A is Toeplitz, q is at most $m + n - 1$, and when $q \le n$, even the general algorithms only require $O(mn^2)$ flops.

For $p = 2$ and $q = mn$, i.e., A is unstructured, the STLN solution coincides with the ordinary TLS solution according to our experiments. Note, however, that the STLN algorithm has a much higher complexity than the SVD-based TLS algorithms [12], [27] and is therefore not recommended for solving unstructured TLS problems.

In contrast to ordinary TLS, LS, and mixed LS-TLS problems [27], the STLN solution can no longer be expressed as the solution of a set of linear equations, and therefore, no closed-form solution for the STLN problem can be derived.

For $p = 2$, the convergence of Algorithm STLN is proven to be equivalent to satisfying the optimality conditions for the minimization problem (2) [21]. In addition, it is shown in [21] that Algorithm STLN is essentially a Gauss–Newton method. Of course, the known limitations of the Gauss–Newton method also apply to Algorithm STLN. For example, we need a small residual for fast convergence.

It should be noted that the differentiable function $\varphi(\alpha, x) = \frac{1}{2} r^H r + \frac{1}{2} \alpha^H D^2 \alpha$ we wish to minimize in (2) for $p = 2$, is not a convex function of (α, x). Therefore, there is no guarantee that a point satisfying the first-order optimality conditions $\nabla_\alpha \varphi = -X_E^H r + D^2\alpha = 0$ and $\nabla_x \varphi = -(A+E)^H r = 0$, is a global minimum. In fact, it could be any stationary point of $\varphi(\alpha, x)$. In general, the Gauss–Newton method will converge to the closest local minimum when the residual $r(\alpha, x)$ is sufficiently small [10]. For STLN problems with $p = 1$ and $p = \infty$, the objective function is not differentiable. Convergence results can be proven (with suitable assumptions) in a way similar to that given in [16] and are currently being studied (see also [21]).

III. STLN for Structured Matrix $[A|b]$

STLN Formulation for Structured $[A|b]$: Frequently, not only A but the right-hand side vector b or $[A|b]$ have a certain structure imposed. For example, in the LS linear prediction problem [14], we need to solve

$$\min_x \|Ax - b\|_2 \tag{7}$$

where $A \in \mathbf{C}^{m\times n}$, $m \ge n$, and either $[A|b]$ or $[b|A]$ is Toeplitz. In this section, it is shown how Algorithm STLN is modified so that it can deal with structured errors in the right-hand side vector b as well [21]. First, a vector β representing possible corrections in selected different elements of b is defined, which is similar to α representing errors in A. When different errors can occur in $l(\le m)$ elements of b, then the corrections applied to these elements are represented by the vector $\beta \in \mathbf{C}^l$. A matrix $F \in \mathbf{R}^{m\times l}$ is introduced so that the corrections applied to b are the same as $F\beta$: The matrix F fans out the elements of β to the corresponding entries in the vector b. The (i, j)th element of F is one if β_j is the correction to b_i; otherwise, it is zero. Initially, E, α, and β are all zero, and the residual $\hat{r} = r = b - Ax$. In general,

$$\begin{aligned} \hat{r} &= \hat{r}(\alpha, \beta, x) \\ &= (b + F\beta) - (A+E)x \\ &= r + F\beta. \end{aligned} \tag{8}$$

From (8), we have

$$\begin{aligned} r &= \hat{r} - F\beta \\ &= b - (A+E)x \end{aligned}$$

which shows that the residual r is divided into two parts $\hat{r}$ and $F\beta$. Therefore, if the structure in E and $F\beta$ allows that $b + F\beta \in Range(A+E)$, then we can require that $\hat{r} = 0$ since $F\beta$ plays the role of r. In this case, the STLN problem that preserves the structures in A and b can be stated as

$$\min_{\hat{r}=0, \alpha, \beta, x} \left\| \begin{bmatrix} \hat{r}(\alpha, \beta, x) \\ D_\alpha \alpha \\ D_\beta \beta \end{bmatrix} \right\|_p \tag{9}$$

for some diagonal weighting matrices D_α and D_β. There are several algorithms that can be used to solve this constrained minimization problem [13]. We use the weighting method for the equality constrained least squares problems, which transforms (9) into an unconstrained problem

$$\min_{\alpha, \beta, x} \left\| \begin{bmatrix} \omega\hat{r}(\alpha, \beta, x) \\ D_\alpha \alpha \\ D_\beta \beta \end{bmatrix} \right\|_p \tag{10}$$

where ω is a large number [2], [29]. It can be shown that for $p = 2$, when ω approaches infinity, the solution of (10) converges to the solution of (9) when the constraint $\hat{r} = 0$ can be satisfied [2], [13], [29]. Thus, by using a large weight ω, we can obtain a good approximation for the solution of (9) by solving the unconstrained problem (10). For possible numerical problems associated with large ω, see [2] and [29]. If these problems occur, a more accurate algorithm for equality constrained problems can be used, such as the one described in [29] based on the generalized SVD (GSVD), when $p = 2$.

The algorithm is summarized in Algorithm STLNB, where the suffix "B" denotes that a structure is imposed on the right-hand side vector b as well. When there is no structure imposed on b, we can simply choose $\beta \in \mathbf{C}^m$ to represent the perturbation on all the elements of b, and $F = I$ accordingly. Therefore, Algorithm STLNB can handle the problems that can be solved by Algorithm STLN, although Algorithm STLN will be more efficient when there is no structure on b.

Algorithm STLNB: **Input**—A structured total least norm problem (9), with specified matrices $[A|b]$, D_α, and D_β, and a tolerance ϵ.

Output—Correction matrix $[E|F\beta]$ and STLNB solution x, such that $[A+E|b+F\beta]$ has the desired structure.

Begin

End

1) Choose a large number ω.
   Set $E = 0$, $\alpha = 0$, $\beta = 0$, compute x from (6) and X_E from x, and set $\hat{r} = b - Ax$.
2) **repeat**
   a)

$$\min_{\Delta\alpha,\Delta\beta,\Delta x} \left\| \hat{M} \begin{pmatrix} \Delta\alpha \\ \Delta\beta \\ \Delta x \end{pmatrix} + \begin{pmatrix} -\omega\hat{r} \\ D_\alpha\alpha \\ D_\beta\beta \end{pmatrix} \right\|_p$$

$$\text{with } \hat{M} = \begin{pmatrix} \omega X_E & -\omega F & \omega(A+E) \\ D_\alpha & 0 & 0 \\ 0 & D_\beta & 0 \end{pmatrix}.$$

   b) Set $x := x + \Delta x$, $\alpha := \alpha + \Delta\alpha$, $\beta := \beta + \Delta\beta$.
   c) Construct E from α and X_E from x. Compute $\hat{r} = (b + F\beta) - (A + E)x$
   until $(\|\Delta x\|, \|\Delta\alpha\|, \|\Delta\beta\| \leq \epsilon)$.

STLNB Algorithm for Toeplitz $[A|b]$*:* For the linear prediction problem where $[A|b]$ (or $[b|A]$) is Toeplitz and all $m+n$ diagonals of $[A|b]_{m\times(n+1)}$ are subject to errors, it is always possible to find a Toeplitz correction matrix $[E|F\beta]$ such that

$$b + F\beta \in Range(A + E).$$

Therefore, we can expect $\hat{r}$ to become zero when the solution is obtained.

The STLNB algorithm is based on the STLN algorithm for structured A and the weighting method for least squares. Therefore, its convergence properties should be analogous to those of the STLN algorithm for structured A, when the weight is large (theoretically when it approaches infinity).

Now, we discuss the case when we need to impose the Toeplitz structure on $[A|b]$. Note that the corrections to b can be represented using the corrections to A except for its first component. If all the elements on the different diagonals of $[A|b]$ are different and subject to error and $\alpha = [\alpha_1 \cdots \alpha_{n+m-1}]^T$, E is Toeplitz with first row $= [\alpha_n \cdots \alpha_1]$ and first column $= [\alpha_n \cdots \alpha_{n+m-1}]^T$, $F = I$, and $\beta = [\beta_1 \cdots \beta_m]^T$, then since $\beta_i = \alpha_{i-1}$, $i = 2, \cdots, m$, we have

$$\begin{aligned} \beta &= \beta_1 e_1 + \hat{F}\alpha \\ &\equiv \beta_1 [1 \quad 0 \quad \cdots \quad 0]^T + \begin{bmatrix} 0_{1\times(m-1)} & 0_{1\times n} \\ I_{(m-1)\times(m-1)} & 0_{(m-1)\times n} \end{bmatrix} \alpha \end{aligned} \tag{11}$$

where I_γ denotes the $\gamma \times \gamma$ identity matrix. In addition, from (11), we have

$$\begin{aligned} &X_E\Delta\alpha - F\Delta\beta + (A+E)\Delta x = \\ &\quad (X_E - \hat{F})\Delta\alpha + (A+E)\Delta x - \Delta\beta_1 e_1 \end{aligned}$$

and Steps 2a) and b) of Algorithm STLNB are simplified to

2a)

$$\min_{\Delta\alpha,\Delta\beta_1,\Delta x} \left\| \hat{M} \begin{pmatrix} \Delta\alpha \\ \Delta\beta_1 \\ \Delta x \end{pmatrix} + \begin{pmatrix} -\omega\hat{r}(\alpha,\beta_1,x) \\ D_\alpha\alpha \\ \beta_1 \end{pmatrix} \right\|_p$$

$$\text{with } \hat{M} = \begin{pmatrix} \omega(X_E - \hat{F}) & -\omega e_1 & \omega(A+E) \\ D_\alpha & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

2b) Set $x := x + \Delta x$, $\alpha := \alpha + \Delta\alpha$, $\beta_1 := \beta_1 + \Delta\beta_1$
$\beta(2:m) := \alpha(1:m-1)$

where $D_\alpha^2 = \text{diag}\,(2, 3, \cdots, n+1, \cdots, n+1, \cdots, 3, 2, 1) \in \mathbf{R}^{(m+n-1)\times(m+n-1)}$, provided all $m+n-1$ elements of A are different and subject to error.

IV. Structured TLN Problems with Multiple Right-Hand Sides

Introduction: Suppose the following set of m linear equations with d right-hand sides is given:

$$A_{m\times n}X \approx B, \quad m > n$$

with

$$\begin{aligned} B_{m\times d} &= [B_{\cdot 1} \cdots B_{\cdot d}], \quad B_{\cdot i} \in \mathbf{C}^m \\ X_{n\times d} &= [X_{\cdot 1} \cdots X_{\cdot d}], \quad X_{\cdot i} \in \mathbf{C}^n \end{aligned} \tag{12}$$

The multiple right-hand side STLN problem can be stated as that of finding E and X such that

$$\min_{E,X} \|[E|-R]\|_p \tag{13}$$

where $R = B - (A+E)X$ and E preserves the structure of A (or another structure can be imposed). Note that the assumption $m > n$ is not a restriction since for $m \leq n$, the set (12) is exactly solvable when A has full rank, and therefore, there is no need to solve a STLN problem in these cases.

For any $m \times d$ matrix $X = [X_{\cdot 1} \cdots X_{\cdot d}]$, where each $X_{\cdot i}$ denotes the ith column of X, we introduce an $md \times 1$ vector vec (X) by stacking the columns of X into one long vector as vec $(X) = [X_{\cdot 1}^T \quad X_{\cdot 2}^T \quad \cdots \quad X_{\cdot d}^T]^T$. The key idea for solving (12) and (13) lies in the fact that (12) can be rewritten as the following set of md linear equations in nd unknowns with one long right-hand side vector: diag $(A, \cdots, A)$ vec $(X) \approx$ vec (B), i.e.,

$$\begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & A \end{pmatrix}_{md\times nd} \begin{pmatrix} X_{\cdot 1} \\ X_{\cdot 2} \\ \vdots \\ X_{\cdot d} \end{pmatrix} \approx \begin{pmatrix} B_{\cdot 1} \\ B_{\cdot 2} \\ \vdots \\ B_{\cdot d} \end{pmatrix}. \tag{14}$$

This implies that the solution matrix X will be found in vectorized form $x = \text{vec}\,(X)$ when (14) is solved by Algorithm STLN if B is unstructured (Algorithm STLNB if B

is structured). The residual matrix $R = B - (A+E)X$ is also obtained in vectorized form vec (R).

STLN Formulation for Structured A with Multiple Right-Hand Side: The multiple right-hand side STLN problem for structured A and unstructured B can be formulated as that of finding α and $x = \text{vec}\,(X)$ such that

$$\min_{\alpha, x} \left\| \begin{bmatrix} r(\alpha, x) \\ D\alpha \end{bmatrix} \right\|_p$$

given a structured A and unstructured B with $A_{m\times n}X \approx B_{m\times d}$, where $r = \text{vec}\,(R) = \text{vec}\,[B - (A+E)X]$. The norm $\|.\|_p$, matrices D and E, and $\alpha \in \mathbf{C}^q$ are as defined for the formulation of the ordinary STLN problem with one right-hand side vector, which is presented in (2).

Construction of the Matrix X_E: In order to compute the STLN solution $x = \text{vec}\,(X)$, we need to construct an $md \times q$ matrix X_E such that

$$\text{diag}\,(E, \cdots, E)x = X_E\alpha. \tag{15}$$

Letting $X_E = [X_{E1}^T \quad X_{E2}^T \quad \cdots \quad X_{Ed}^T]^T$, (15) gives $EX_{\cdot i} = X_{Ei}\alpha$ for $1 \le i \le d$. Because X_{Ei} is completely determined by the structure of E and E itself is independent of the column index i of X, all X_{Ei}'s have the same structure. The only difference is the values of the matrix elements. Namely, X_{E1} is filled with the n elements of $X_{\cdot 1}$, X_{E2} with the n elements of $X_{\cdot 2}$, etc. For example, consider the example given in (3) in Section II, and assume that there are two right-hand side vectors. Then, we have

$$X_E = \begin{pmatrix} X_{E1} \\ X_{E2} \end{pmatrix} = \begin{pmatrix} x_{31} & x_{21} & x_{11} & 0 & 0 & 0 \\ 0 & x_{31} & x_{21} & x_{11} & 0 & 0 \\ 0 & 0 & x_{31} & x_{21} & x_{11} & 0 \\ 0 & 0 & 0 & x_{31} & x_{21} & x_{11} \\ x_{32} & x_{22} & x_{12} & 0 & 0 & 0 \\ 0 & x_{32} & x_{22} & x_{12} & 0 & 0 \\ 0 & 0 & x_{32} & x_{22} & x_{12} & 0 \\ 0 & 0 & 0 & x_{32} & x_{22} & x_{12} \end{pmatrix}$$

where

$$X = [X_{\cdot 1} \quad X_{\cdot 2}] = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{pmatrix}.$$

Using this relationship and linearizing $r(\alpha + \Delta\alpha, x + \Delta x)$ yields a linearized minimization problem similar to (5):

$$\min_{\Delta\alpha, \Delta x} \left\| M \begin{pmatrix} \Delta\alpha \\ \Delta x \end{pmatrix} + \begin{pmatrix} -r \\ D\alpha \end{pmatrix} \right\|_p \tag{16}$$

where

$$M = \begin{bmatrix} X_E & \text{diag}\,(A+E, \cdots, A+E)_{md\times nd} \\ D & 0 \end{bmatrix},$$
$$\Delta x = \text{vec}\,(\Delta X),$$
$$X_E = [X_{E1}^T \quad \cdots \quad X_{Ed}^T]^T.$$

The initial solution for x is easily obtained by solving d minimization problems:

$$\min_{X_{\cdot i}} \|B_{\cdot i} - AX_{\cdot i}\|_p \quad i = 1, \cdots, d.$$

In particular, for $p = 2$, each $X_{\cdot i}$ is initialized as $(A^HA)^{-1}A^HB_{\cdot i}$, which requires only one QR decomposition of A. By exploiting the structure of the corresponding matrix M, the QR factorization can be performed in an efficient way as well when solving (16) with $p = 2$.

STLN for Structured $[A|B]$ with Multiple Right-Hand Side: Similarly, the multiple right-hand side STLN problem for structured $[A|B]$ can be formulated as that of finding $\alpha \in \mathbf{C}^q (q \le mn)$, $\beta \in \mathbf{C}^l (l \le md)$, and $x = \text{vec}\,(X)$ such that

$$\min_{\hat{r}=0, \alpha, \beta, x} \left\| \begin{bmatrix} \hat{r}(\alpha, \beta, x) \\ D_\alpha\alpha \\ D_\beta\beta \end{bmatrix} \right\|_p \tag{17}$$

given $A_{m\times n}X \approx B_{m\times d}$, where $[A|B]$ is structured, $\hat{r} = \text{vec}\,(\hat{R}) = \text{vec}\,(R) + F\beta = r + F\beta = b + F\beta - \text{diag}\,(A + E, \cdots, A + E)x$, $x = \text{vec}\,(X)$, $b = \text{vec}\,(B)$ and $F\beta = [F_1^T \quad \cdots \quad F_d^T]^T\beta$. Each $F_i \in R^{m\times l}$, $1 \le i \le d$, is defined so that the corrections applied to each $B_{\cdot i}$ are given by $F_i\beta$.

The multiple right-hand side STLNB solution is computed using algorithm STLNB, provided $\hat{r}$ and F, as defined above, and X_E, as defined in (15), are used, and $A+E$ is replaced by diag $(A+E, \cdots, A+E)$ in Step 2(a) of Algorithm STLNB.

STLN for Toeplitz $[A|B]$ with Multiple Right-Hand Side: Assume that $[A|B]$ is Toeplitz, and $[E|\,F_1\beta \quad \cdots \quad F_d\beta]$ needs to preserve the same structure. Representing corrections in A by α_i's and in B by β_i's such that

$$E = \begin{pmatrix} \alpha_n & \alpha_{n-1} & \cdots & \alpha_1 \\ \alpha_{n+1} & \alpha_n & \cdots & \alpha_2 \\ \vdots & \ddots & \ddots & \vdots \\ \alpha_{n+m-1} & \alpha_{n+m-2} & \cdots & \alpha_m \end{pmatrix}$$

and

$$[F_1\beta \quad \cdots \quad F_d\beta] = \begin{pmatrix} \beta_d & \beta_{d-1} & \cdots & \beta_1 \\ \beta_{d+1} & \beta_d & \cdots & \beta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{d+m-1} & \beta_{d+m-2} & \cdots & \beta_m \end{pmatrix}$$

with $\alpha = [\alpha_1 \quad \cdots \quad \alpha_{n+m-1}]^T$ and $\beta = [\beta_1 \quad \cdots \quad \beta_{d+m-1}]^T = [\beta_1 \quad \cdots \quad \beta_d\,\alpha_1 \quad \cdots \quad \alpha_{m-1}]^T$, we have

$$F_i\beta = J_i\hat{\beta} + \hat{F}_i\alpha \quad \hat{\beta} = [\beta_1 \quad \cdots \quad \beta_d]^T$$

where

$$\begin{cases} J_i = \begin{pmatrix} 0 & I_i \\ 0 & 0 \end{pmatrix} \text{ and } \hat{F}_i = \begin{pmatrix} 0 & 0 \\ I_{m-i} & 0 \end{pmatrix} \text{ if } i < m \\ J_i = (0 \quad I_m \quad 0) \text{ and } \hat{F}_i = 0 \quad \text{if} \quad i \ge m \text{ (when } d \ge m) \\ J_i \text{ an } m \times d \text{ matrix and } F_i \text{ an } m \times (m+n-1) \text{ matrix.} \end{cases}$$

Using these notations, Steps 2(a) and (b) of the Algorithm STLNB are simplified to

2a)

$$\min_{\Delta\alpha,\,\Delta\beta,\,\Delta x} \left\| \hat{M} \begin{pmatrix} \Delta\alpha \\ \Delta\hat{\beta} \\ \Delta x \end{pmatrix} + \begin{bmatrix} -\omega\hat{r} \\ D_\alpha \alpha \\ D_{\hat{\beta}}\beta \end{bmatrix} \right\|_p$$

with

$$\hat{M} = \begin{bmatrix} \omega(X_E - \hat{F}) & -\omega J & \omega\,\mathrm{diag}(A+E, \cdots, A+E) \\ D_\alpha & 0 & 0 \\ 0 & D_{\hat{\beta}} & 0 \end{bmatrix}.$$

2b) Set $x := x + \Delta x$, $\alpha := \alpha + \Delta\alpha$, $\beta(1:d) := \beta(1:d) + \Delta\hat{\beta}$, $\beta(d+1:d+m-1) := \alpha(1:m-1)$

with

$$\begin{pmatrix} D_{\hat{\beta}}^2 & 0 \\ 0 & D_\alpha^2 \end{pmatrix} = \mathrm{diag}\,(1, 2, \cdots, p, \cdots, p, \cdots, 2, 1),$$

$$D_\alpha \in \boldsymbol{R}^{(m+n-1)\times(m+n-1)},$$

$$D_{\hat{\beta}} \in \boldsymbol{R}^{d\times d},$$

$$p = \min\,(m, n+d),$$

$$J = [J_1^T \ \cdots \ J_d^T]^T$$

and

$$\hat{F} = [\hat{F}_1^T \ \cdots \ \hat{F}_d^T]^T.$$

Application: As an important byproduct, Algorithm STLNB can be used to compute the best rank r structure-preserving approximation C_{stln} to a given matrix $C_{m\times n}$, $m > n$, of rank r_C in the sense that $\|C - C_{stln}\|_p$ is minimized, $0 < \mathrm{rank}\,(C_{stln}) \le r < r_C$, and C_{stln} has the same structure as C. This can be accomplished by the following procedure:

- Set $C = [A_{m\times r} | B_{m\times(n-r)}]$ with rank $(A) = r$.
- Compute the STLN solution α_{stln}, β_{stln} of the corresponding STLNB minimization problem (17) with $n - r$ right-hand sides.
- Construct E_{stln} from α_{stln}.
- Then, $C_{stln} = [A + E_{stln} | B + [F_1\beta_{stln} \ \cdots \ F_{n-r}\beta_{stln}]]$.

This procedure can be used as signal enhancement method in exponential data modeling problems [4] and for model reduction of multi-input multi-output systems such that some measure of the difference between the reduced order model and the original model is minimized (see e.g., [11]). Preliminary promising results of the former application are reported in [3]. For the latter application, time-domain algorithms based on the structured TLS method [6] and [8], which are optimal with respect to a finite-horizon L_2-norm, have been developed recently for single-input single-output systems and are discussed in [9] and [23]. Some of these algorithms have been transformed to the z domain, thereby enabling computation of the optimal reduced model with respect to the L_2 norm (over an infinite horizon) and the Hilbert–Schmidt–Hankel norm [7], [22]. The most salient shortcoming of these STLS-type approximation algorithms is the fact that they cannot deal with multi-input multi-output systems because the STLS method cannot be easily extended to reduce the rank of a given structured matrix by more than one. Since the STLN method can solve this problem as shown above, it is expected that this opens the way to develop STLN-type approximation algorithms for model reduction of multi-input multi-output systems.

V. Optimality of the STLN Solution in Model-Reduction Problems

We now discuss modeling problems, which approximate a given data sequence by the impulse response of a finite-dimensional linear time-invariant (LTI) system of a given order, such that the weighted sum of squared residuals is minimized. This problem can be solved by the STLN algorithm. We show that in the examples we tested, the STLN method converges to the optimal solution.

Consider the problem of approximating a given data sequence $z_k \in \mathbf{C}^{m+s-1}$ by $\tilde{z}_k \in \mathbf{C}^{m+s-1}$ to minimize

$$\sum_{k=1}^{m+s-1} [(z_k - \tilde{z}_k)w_k]^2 \quad \text{subject to} \quad \tilde{C}\begin{pmatrix} X \\ -I_d \end{pmatrix} = 0 \qquad (18)$$

where $\tilde{C}$ is an $m \times s$ Hankel or Toeplitz matrix, and $m \ge s$, which has been constructed from the elements $\tilde{z}_k$, and w_k are appropriate weights. The constraint imposing that $\tilde{C}$ has at most rank $s - d$ ensures that $\tilde{z}_k$ is the impulse response of a finite-dimensional linear system of order $s - d$ at most. If the sequence z_k is itself the impulse response of a higher dimensional system, our problem corresponds to *model reduction.* If the sequence z_k is a given data sequence that is not an impulse response but, for instance, a noise-corrupted one, one might consider this problem as a noisy realization problem. If, additionally, all $w_k = 1$, we get model reduction in the L_2 norm.

If the weights w_k account for the repetition of elements $\tilde{z}_k$ in the matrix $\tilde{C}$, then

$$\begin{pmatrix} D_{\hat{\beta}}^2 & 0 \\ 0 & D_\alpha^2 \end{pmatrix} = \mathrm{diag}\,(1, 2, 3, \cdots, s, \cdots, s, \cdots, 3, 2, 1)$$

and (18) reduces to a STLNB minimization problem with d right-hand sides, where

$$[\hat{\beta}^T, \alpha^T] = [\tilde{z}_1 - z_1 \ \cdots \ \tilde{z}_{m+s-1} - z_{m+s-1}]$$

and $m + s - 1$ is the number of different elements in $C = [A_{m\times n} | B_{m\times d}]$, $s = n + d$, which are all subject to error. The full-rank Toeplitz matrix C has its first column $= [z_s \ \cdots \ z_{m+s-1}]^T$ and first row $= [z_s \ \cdots \ z_1]$ and needs to be reduced to a rank n matrix $\tilde{C}$, which is similarly constructed as C with the elements $\tilde{z}_k$.

As proven in [8], there is one special case for which the global optimum of (18) can be found explicitly, i.e.

Approximate a given data sequence z_k by a first-order LTI impulse response, which can be parametrized as $\tilde{z}_k = \gamma\eta^{k-1}$.

TABLE I
APPROXIMATION OF THE RANK 4 TOEPLITZ MATRIX $C_{5\times4}$, DESCRIBED IN EXAMPLE 5.2, BY A SIMILARLY STRUCTURED RANK 1 TOEPLITZ MATRIX $\tilde{C}$ SUCH THAT $\|C-\tilde{C}\|_F$ IS MINIMIZED. THE ALGORITHM STLNB (SOLVING THE EQUALITY CONSTRAINED LS PROBLEM BY MEANS OF THE GSVD BASED METHOD DESCRIBED IN [29]) CONVERGES WITHIN 11 ITERATIONS. CADZOW'S APPROACH CONVERGES WITHIN 85 ITERATIONS

Method	ν	$\|C-\tilde{C}\|_F$
rooting polynomial	0.2602566142908349	0.1030824769957293
algorithm STLNB	0.2602566142908349	0.1030824769957293
Cadzow's approach	0.2602455147443434	0.1030825010827233

We then obtain a minimization problem:

$$\min_{\gamma,\eta}\sum_{k=1}^{m+s-1}[(z_k-\gamma\eta^{k-1})w_k]^2=\min_{\gamma,\eta}g(\gamma,\eta). \tag{19}$$

Setting the derivatives with respect to γ and η to zero, we obtain

$$\frac{\partial g}{\partial\gamma}=0\Rightarrow\sum_{k=1}^{m+s-1}(z_k-\gamma\eta^{k-1})w_k^2(-\eta^{k-1})=0,$$
$$\frac{\partial g}{\partial\eta}=0\Rightarrow\sum_{k=2}^{m+s-1}(z_k-\gamma\eta^{k-1})w_k^2[-(k-1)\eta^{k-2}]=0.$$

Eliminating γ from the first equation and substituting it in the second, we get the polynomial in η :

$$\left[\sum_{k=2}^{m+s-1}(k-1)z_kw_k^2\eta^{k-2}\right]\left[\sum_{k=1}^{m+s-1}w_k^2\eta^{2k-2}\right]-\left[\sum_{k=1}^{m+s-1}z_kw_k^2\eta^{k-1}\right]\left[\sum_{k=2}^{m+s-1}(k-1)w_k^2\eta^{2k-3}\right]=0. \tag{20}$$

This is a polynomial of degree $3(m+s-1)-5$ in η. One has to select the root that gives the minimum in (19).

Example 5.1: Consider the same data sequence $z=[6\quad5\quad4\quad3\quad2\quad1]^T$, as given in [8], which can be modeled as the impulse response of the following second-order LTI system:

$$w_{k+1}=\begin{pmatrix}1&-1\\0&1\end{pmatrix}w_k;$$
$$z_k=[1\quad0]\,w_k\quad\text{with}\quad w_0=\begin{pmatrix}6\\1\end{pmatrix}.$$

We arrange the data z in the 5×2 Toeplitz matrix

$$C=\begin{pmatrix}z_2&z_1\\z_3&z_2\\z_4&z_3\\z_5&z_4\\z_6&z_5\end{pmatrix}=\begin{pmatrix}5&6\\4&5\\3&4\\2&3\\1&2\end{pmatrix}$$

and find the best rank 1 structure-preserving approximation

$$\tilde{C}=\begin{pmatrix}\tilde{z}_2&\tilde{z}_1\\\tilde{z}_3&\tilde{z}_2\\\tilde{z}_4&\tilde{z}_3\\\tilde{z}_5&\tilde{z}_4\\\tilde{z}_6&\tilde{z}_5\end{pmatrix}$$

to C by solving the corresponding STLN problem with weights $(w_1^2, w_2^2, \cdots, w_6^2)=(1,2,2,2,2,1)$, $m=5$, $n=1$, $d=1$. The true minimum $\eta=0.762\,923\,015\,074\,321\,8$ of (18), which is obtained by finding the roots of (20) using MATLAB version 4.2a, was also obtained after 15 iterations of the STLNB algorithm, up to 14 digits, from each ratio $\tilde{z}_k/\tilde{z}_{k-1}\,k=2,\cdots,6$. This shows that the algorithm STLNB converges to the optimal solution, for this example, just as the structured TLS method [6]. On the other hand, the suboptimal methods such as Cadzow's approach [4] do not converge to the actual solution (as shown in [8]).

Example 5.2: Consider the following fourth-order LTI system:

$$w_{k+1}=\text{diag}\,(0.4,0.3,0.2,0.1)w_k;$$
$$z_k=[1\quad1\quad1\quad1]w_k$$

with

$$w_0=[1\quad1\quad1\quad1]^T.$$

Arranging the first eight samples z_k of the impulse response in a 5×4 Toeplitz matrix C with first row $=[0.1\quad0.3\quad1\quad4]$ and first column $=[0.1\quad0.0354\quad0.013\quad0.004\,89\quad0.001\,87]^T$ and computing the best rank 1 structure-preserving approximation $\tilde{C}$ to C by solving the corresponding STLN problem with three right-hand sides ($d=3$), weights $(w_1^2, w_2^2, \cdots, w_8^2)=(1,2,3,4,4,3,2,1)$, $m=5$ and $n=1$, we again reach the known optimal solution for η within 13 iterations of the STLNB algorithm, as shown in Table I. In addition, it is shown that Cadzow's approach converges to a suboptimal solution. Since the Structured TLS method [6] cannot reduce the rank of C by more than one, this method cannot be considered here. Similar results have been obtained for various other sequences z_k and other weights w_k (including the unweighted case where all $w_k=1$) for the problem (19).

VI. COMPUTATIONAL RESULTS: COMPARISON OF STLN WITH LS AND TLS IN PARAMETER ESTIMATION PROBLEMS

The algorithms STLN and STLNB have been implemented in MATLAB in order to investigate their performance in problems where structured matrices are involved. In particular,

TABLE II
RELATIVE ACCURACY OF THE (a) LP SOLUTION, (b) DAMPING FACTOR, (c) FREQUENCY IN THE FOLLOWING LP PROBLEM: $m = 42$, $n = 8$, $d = 1$, $z_t = \sum_{k=1}^{4} e^{(-\alpha_k + 2\pi\sqrt{-1} f_k)t}$, $1 \le t \le 50$, THE FIRST COLUMN OF A_c IS $[z_{42} \ \cdots \ z_2 \ z_1]^T$ AND $b_c = [z_{50} \ \cdots \ z_9]^T$, SO THAT $[A_c|b_c]$ IS TOEPLITZ. THE EIGHT PAIRS (d_k, f_k) ARE (0.1, 0.5), (0.2, 0.4), (0.3, 0.3), (0.35, 0.1), (0.4, 0.2), (0.5, 0.45), (0.05, 0.25), (0.45, 0.05)

σ_ν	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}	δd_{ls}^{rel}	δd_{tls}^{rel}	δd_{stln}^{rel}	δf_{ls}^{rel}	δf_{tls}^{rel}	δf_{stln}^{rel}
e-10	7.12e-8	7.12e-8	2.04e-9	1.34e-7	1.34e-7	4.02e-9	1.98e-8	1.98e-8	6.8e-10
e-9	6.61e-7	6.61e-7	1.87e-8	1.18e-6	1.18e-6	3.87e-8	1.98e-7	1.98e-7	6.06e-9
e-8	7.19e-6	7.19e-6	1.98e-7	1.30e-5	1.30e-5	3.54e-7	2.05e-6	2.05e-6	6.40e-8
e-7	7.12e-5	7.12e-5	1.97e-6	1.35e-4	1.35e-4	3.74e-6	1.99e-5	1.99e-5	6.87e-7
e-6	6.95e-4	6.94e-4	1.91e-5	1.27e-3	1.27e-3	3.72e-5	2.00e-4	2.01e-4	6.63e-6
e-5	8.03e-3	6.53e-3	2.12e-4	1.29e-2	1.15e-2	3.86e-4	2.32e-3	1.95e-3	6.94e-5
e-4	3.66e-1	8.29e-2	1.86e-3	9.27e-1	9.62e-2	3.53e-3	6.19e-1	3.06e-1	6.54e-4
	(a)			(b)			(c)		

we here compare the accuracy of these algorithms with that of the currently used LS and TLS methods in several signal processing and system identification problems in order to show the gain in accuracy one can expect by exploiting the structure of the problem. Other computational tests are reported in [21]. For the effect on the algorithm behavior of structure, initial choice of x, the magnitude of the minimum norm, and the choice of the L_p norm, see also [21].

Test Procedure: We start from $m + n - 1$ data samples z_i arranged in an $m \times n$ Toeplitz matrix A_c and set up the zero residual set of equations $A_c X_c = B_c$ by either choosing a specific exact solution X_c or right-hand side matrix B_c. Random errors that are normally distributed with mean zero and variance σ_ν^2 are then added to the data z_i and B_c, generating the perturbed set $AX \approx B$. We denote the vectorized LS, TLS, and STLN solution of this perturbed set by x_{ls}, x_{tls}, x_{stln}, respectively. The relative errors $\delta_\alpha^{rel} \equiv \|x_\alpha - x_c\|/\|x_c\|$ with $\alpha = ls$, tls, or $stln$ are averaged over 100 runs per σ_ν. These tests are repeated for progressively larger σ_ν. If not stated otherwise, the L_2 norm STLN solution is computed in 10 iterations, and the matrix M always has full column rank.

Test Problems: We first consider a *linear prediction* (LP) problem, which is described in Table II. In addition to the relative accuracy of the LP solution vector x, we show the relative accuracy of damping factor and frequency estimates obtained by finding the roots of the computed LP polynomial in each run, as in Prony's method. Observe that STLN improves the accuracy of the TLS estimates with a factor 30 to 40 for noise standard deviations up to 10^{-5}, whereas the differences in accuracy between TLS and LS remain negligible! For $\sigma_\nu = 10^{-4}$, the frequency accuracy improves by a factor of more than 400.

Algorithm STLNB can also be used for computing the rank r Toeplitz approximation of any $m \times n$ Toeplitz matrix C, as described at the end of Section IV and shown below. The above LP problem is extended to a problem with multiple right-hand sides, as described below. Consider the Toeplitz matrix $C_c = [A_c|B_c]$, where A_c is the same matrix as described in Table II ($m = 42, n = 8$), but now, four extra columns B_c ($d = 4$) are added, keeping the Toeplitz

TABLE III
RELATIVE ACCURACY OF THE LS, TLS, AND STLN SOLUTION WHEN SOLVING THE NOISY SETS $AX \approx B$ WHERE $[A|B]$ IS TOEPLITZ. A_c AND THE FIRST COLUMN OF B_c ARE DESCRIBED BY A_c AND b_c IN TABLE II. THE FIRST ROW OF $B_c = [z_{50} \ \ z_{51} \ \ z_{52} \ \ z_{53}]$

σ_ν	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}
e-10	6.43e-8	6.43e-8	4.38e-9
e-8	7.35e-6	7.35e-6	4.47e-7
e-6	6.20e-4	6.15e-4	4.48e-5

structure. Perturb C_c to C, as described in the test procedure, and compute the best rank 8 matrix approximation for C using the LS, TLS, and STLN method. Only the STLN method is able to preserve the Toeplitz structure. The results are described in Table III. Observe that STLN improves the solution accuracy by a factor of 15. In all test cases considered here, the d smallest singular values of the STLN approximation $[A + E|B + (F_1\beta \ \cdots \ F_d\beta)]$ were smaller than 10^{-14} after 10 iterations. We also computed the average Frobenius norm corr_{ls}, corr_{tls}, and corr_{stln} of the corresponding LS, TLS, and STLN correction matrices $[E_\gamma|F_\gamma]$ computed by means of method γ such that $(A + E_\gamma)X_\gamma = B + F_\gamma$, where $\gamma =$ LS, TLS, or STLN. For $\sigma_\nu = 10^{-8}$, we obtained $\text{corr}_{ls} = 6.20\times10^{-6}$, $\text{corr}_{tls} = 1.24\times10^{-7}$, and $\text{corr}_{stln} = 1.89\times10^{-7}$. The average Frobenius norm of the true perturbation matrix was $\text{corr}_{true} = 2.22 \times 10^{-7}$. This is a typical result. As can easily be proven, corr_{tls} is always minimal; corr_{stln} is only slightly larger and always less than corr_{true}, whereas corr_{ls} is considerably larger than all other values.

Next, we explore the use of STLN in *time-domain system identification.* If the process can be modeled as an LTI, causal, finite-dimensional system with zero-initial state, then an *impulse response model* may be used:

$$y(t) = \sum_{k=0}^{n} h(k)u(t-k). \tag{21}$$

The system is identified if its impulse response $h(k)$, $k = 0, \cdots, n-1$ can be estimated from observations of the inputs $u(t)$ and outputs $y(t)$ over a certain interval of time $t = 1-n, \cdots, m-1$. This so-called *deconvolution* problem

TABLE IV
RELATIVE ACCURACY OF THE IMPULSE RESPONSE IN THE FOLLOWING DECONVOLUTION PROBLEM: $m = 30$, $n = 20$, $d = 1$; $h = x_c = [1.9\ 3.3\ 4.4\ 5.4\ 5.9\ 6.2\ 6.3\ 6.1\ 5.8\ 5.6\ 5.3\ 5.0\ 4.85\ 4.6\ 4.0\ 3.4\ 1.8\ 1.0\ 0.2\ 0.0]^T$ IS THE EXACT IMPULSE RESPONSE. THE INPUTS $u(t)$, $t = -19, \cdots, 29$ ARE NORMALLY DISTRIBUTED ZERO-MEAN NOISE OF UNIT VARIANCE: (a) $u(t) \neq 0\ \forall t$, (b) $u(t) = 0\ \forall t < 0$

σ_ν	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}
e-10	1.6e-10	1.6e-10	5.5e-11	2.7e-10	2.7e-10	1.4e-10
e-8	1.76e-8	1.76e-8	6.41e-9	3.10e-8	3.10e-8	1.67e-8
e-6	1.81e-6	1.81e-6	5.85e-7	3.28e-6	3.28e-6	1.66e-6
e-4	1.51e-4	1.51e-4	5.98e-5	3.32e-4	3.32e-4	1.70e-4
e-3	1.77e-3	1.77e-3	6.22e-4	2.98e-3	2.98e-3	1.50e-3
e-2	1.71e-2	1.71e-2	5.67e-3	3.04e-2	3.04e-2	1.56e-2
e-1	1.85e-1	1.83e-1	6.11e-2	3.34e-1	3.81e-1	1.66e-1
	(a)			(b)		

is essentially reduced to a problem of solving a set of linear equations $Ax \approx b$ by writing (21) for $t = 0, \cdots, m-1$, i.e.,

$$\begin{bmatrix} u(0) & u(-1) & \cdots & u(1-n) \\ u(1) & u(0) & \cdots & u(2-n) \\ \vdots & \vdots & \ddots & \vdots \\ u(m-1) & u(m-2) & \cdots & u(m-n) \end{bmatrix} \begin{bmatrix} h(0) \\ \vdots \\ (n-1) \end{bmatrix} \approx \begin{bmatrix} y(0) \\ \vdots \\ y(m-1) \end{bmatrix}. \tag{22}$$

The STLN algorithm takes into account the Toeplitz structure of the data matrix A, and this improves the accuracy of the TLS and LS impulse response estimates by a factor 3, as shown in Table IV(a). If $u(t)$ is zero for $t < 0$, then the strictly upper triangular part of A is zero. The STLN is able to take also this information into account, thereby improving the TLS and LS estimates by a factor of 2 (see Table IV(b)).

To illustrate the identification of multivariable systems with multiple inputs and/or outputs, we consider the same system as described in Table IV with one input but three outputs $y_1(t)$, $y_2(t)$, and $y_3(t)$. The corresponding impulse responses $h_1(t)$, $h_2(t)$, and $h_3(t)$ can be estimated by solving the corresponding deconvolution problem with three unstructured right-hand sides $B = [y_1 \quad y_2 \quad y_3]$, given, respectively, by the three outputs. In this example, the number of STLN iterations is not kept fixed to 10 but is made adaptive according to certain criteria, as specified in Table V. In this table, the results are displayed, and they show that STLN improves the accuracy of the TLS and LS estimates by a factor of 3 or more. In addition, observe that the STLN algorithm converges very fast, especially when the perturbations are small. In all tests considered, the three smallest singular values of the STLN approximation $[A + E|B + (F_1\beta \quad F_2\beta \quad F_3\beta)]$ were reduced to order 10^{-15}, implying that all data matrices $C = [A|B]$ converged to a rank-deficient matrix of rank 20 within machine precision while preserving the structure of A. Moreover, the corresponding STLN correction matrix is only slightly larger than the TLS correction matrix, whereas the LS correction matrix is considerably larger than all other values.

Another frequently used model to describe a system behavior is the *transfer function* model, which essentially expresses that the inputs $u(t)$ and outputs $y(t)$ fed to the system satisfy a linear difference equation. In polynomial form, we have

$$\begin{aligned}(1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}) y(t) \\ = (b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}) u(t)\end{aligned} \tag{23}$$

with q^{-1} the backward shift operator. Since $q^{-1} y(t) = y(t-1)$, (23) reduces to

$$\begin{aligned}y(t) + a_1 y(t-1) + \cdots + a_{n_a} y(t-n_a) \\ = b_1 u(t-1) + \cdots + b_{n_b} u(t-n_b).\end{aligned} \tag{24}$$

Writing out (24) for $t = t_0, \cdots, t_0 + m - 1$ with t_0 a given offset, one obtains an overdetermined $m \times (n_a + n_b)$ set of equations: $y = [Y_1|U_1]x$ with $y = [y(t_0) \ \cdots \ y(t_0+m-1)]^T$, $x = [a_1 \ \cdots \ a_{n_a}\ b_1 \ \cdots \ b_{n_b}]^T$, and the tth row of $[Y_1|U_1]$ is $[-y(t-1) \ \cdots \ -y(t-n_a)\, u(t-1) \ \cdots \ u(t-n_b)]$. This matrix is structured as a [Toeplitz|Toeplitz] matrix. For simplicity, we consider only models with a pure Toeplitz data matrix, i.e.,

a) models with $n_a = 0$ and modeling equations: $y = U_1 x$..
b) models with $n_b = 0$ and modeling equations: $y = Y_2 \hat{x}$. $\hat{x} = [-a_{n_a} \ \cdots \ -a_1]^T$. The tth row of Y_2 is $[y(t-n_a) \ \cdots \ y(t-1)]$ such that $[Y_2|y]$ is Toeplitz (the same as the LP equations)
c) models with $n_b = 1$ and modeling equations: $[y|-Y_1]\check{x} = u$ with $u = [u(t_0-1) \ \cdots \ u(t_0+m-2)]^T$ and $\check{x} = [1 \quad a_1 \ \cdots \ a_{n_a}]^T / b_1$.

In Table VI, an example of each class is considered. The beneficial effect of taking the structure into account is clearly shown at any signal-to-noise ratio (SNR). Indeed, for all noise variances considered, the STLN estimates of the coefficients a_i, b_i are two to four times more accurate than the corresponding TLS estimates. Note also that the gain in accuracy obtained by ordinary TLS compared with LS is negligible for high SNR. Only at very low SNR, the gain in accuracy can improve up to a factor 8.

VII. CONCLUSIONS AND FUTURE WORK

A recently presented new method [20] and [21] related to total least squares (TLS) known as structured total least norm (STLN) for solving overdetermined sets of linear equations $Ax \approx b$ has been summarized, where A or $[A|b]$ is structured and affected by noise. The main new features of this approach are that it preserves the problem structure and permits the minimization of error in different norms as well, which is important in certain problems to improve robustness of the solution.

The STLN problem is generalized to structured sets $AX \approx B$ with multiple right-hand sides. It is shown how this problem can be converted to an ordinary STLN problem with one right-hand side so that the same algorithms can be used. In particular, the multiple right-hand-side STLN method allows us to compute the best structure-preserving approximation of any lower rank to a given structured matrix. In addition, we have shown by test examples that the STLN algorithm

TABLE V

RELATIVE ACCURACY OF THE IMPULSE RESPONSES IN THE FOLLOWING DECONVOLUTION PROBLEM: $m = 30$, $n = 20$, $d = 3$; $h_1 = X_{.1}$ EQUALS $\underline{h}$ AS DESCRIBED IN TABLE IV; $h_2 = X_{.2} = [0.29\ 0.43\ 0.64\ 0.74\ 0.89\ 0.92\ 0.93\ 0.91\ 0.78\ 0.76\ 0.73\ 0.60\ 0.585\ 0.56\ 0.50\ 0.44\ 0.28\ 0.10\ 0.02\ 0.0]^T$ AND $h_3 = X_{.3} = [0.09\ 0.23\ 0.34\ 0.44\ 0.49\ 0.52\ 0.58\ 0.56\ 0.53\ 0.51\ 0.48\ 0.45\ 0.40\ 0.36\ 0.30\ 0.24\ 0.08\ 0.010\ 0.002\ 0.0]^T$ ARE THE EXACT IMPULSE RESPONSES. THE INPUTS $u(t)$, $t = -19, \cdots, 29$, ARE NORMALLY DISTRIBUTED ZERO-MEAN NOISE OF UNIT VARIANCE. *it* IS THE AVERAGE NUMBER OF STLN ITERATIONS USED (ITERATIONS ARE STOPPED WHEN $\|\Delta\|_2 = \|[\Delta\alpha^T\ \Delta x^T]^T\|_2 \leq \text{SIZE}(\Delta)10^{-15}$ OR THE MAXIMAL NUMBER OF ITERATIONS (SET TO 50) IS REACHED). corr_γ DENOTES THE AVERAGE FROBENIUS NORM OF THE CORRECTION MATRIX $[E_\gamma|F_\gamma]$ COMPUTED BY MEANS OF METHOD γ SUCH THAT $(A + E_\gamma)X_\gamma = B + F_\gamma$ WHERE $\gamma = ls$, *tls*, *stln* OR *true*

σ_ν	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}	*it*	corr_{ls}	corr_{tls}	corr_{stln}	corr_{true}
e-10	1.63e-10	1.63e-10	5.64e-11	2	1.32e-8	5.39e-10	1.04e-9	2.64e-9
e-8	1.74e-8	1.74e-8	5.30e-9	2	1.31e-6	5.45e-8	1.05e-7	2.59e-7
e-6	1.68e-6	1.68e-6	5.23e-7	3	1.46e-4	5.68e-6	1.06e-5	2.61e-5
e-4	1.64e-4	1.64e-4	5.43e-5	4	1.36e-2	5.46e-4	1.06e-3	2.62e-3
5e-4	8.89e-4	8.88e-4	2.66e-4	5	7.16e-2	2.76e-3	5.38e-3	1.31e-2
e-3	1.65e-3	1.65e-3	5.58e-4	5.34	1.45e-1	5.59e-3	1.07e-2	2.64e-2
5e-3	7.90e-3	7.94e-3	2.60e-3	6.92	6.68e-1	2.64e-2	5.06e-2	1.28e-1
e-2	1.65e-2	1.64e-2	5.49e-3	8.01	1.50e+0	5.73e-2	1.06e-1	2.61e-1
5e-2	7.71e-2	7.59e-2	2.62e-2	12.89	7.06e+0	2.77e-1	5.27e-1	1.29e+0
e-1	1.57e-1	1.77e-1	5.50e-2	17.22	1.40e+1	5.49e-1	1.04e+0	2.62e+0

TABLE VI

RELATIVE ACCURACY OF THE PARAMETERS OF THE FOLLOWING TRANSFER FUNCTION MODELS: (a) $\frac{1}{95}y(t) = (q^{-1} + .99q^{-2} + .98q^{-3} + .97q^{-4} + .96q^{-5} + .92q^{-6} + .87q^{-7})u(t)$, $m = 30$, $n = n_b = 7$, $d = 1$ AND $u(t) = \sum_{k=1}^{7} a_k \cos[2\pi f_k(t-1)\Delta]$, $t = 1, \cdots, 36$ AND $\Delta = .01$. THE 7 PAIRS (a_k, f_k) ARE $(e^{-0.1}, 42)$, $(e^{-0.02}, 52)$, $(e^{-0.3}, 52.1)$, $(e^{-0.4}, 60)$, $(e^{-0.5}, 90)$, $(.1e^{-0.8}, 92)$, $(.01e^{-0.03}, 85)$; (b) $(1 + .98q^{-1} + .95q^{-2} + .92q^{-3} + .9q^{-4} + .85q^{-5} + .8q^{-6} + .76q^{-7})y(t) = 0$, $m = 60$, $n = n_a = 7$, $d = 1$ AND $y(t)$ NORMALLY DISTRIBUTED ZERO-MEAN NOISE OF UNIT VARIANCE; (c) $(1 + .99q^{-1} + .98q^{-2} + .97q^{-3} + .96q^{-4} + .92q^{-5} + .9q^{-6})y(t) = .063u(t)$, $m = 30$, $n + n_a + n_b = 6 + 1 = 7$, $d = 1$ AND $u(t)$ NORMALLY DISTRIBUTED ZERO-MEAN NOISE OF UNIT VARIANCE

σ_ν	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}	δx_{ls}^{rel}	δx_{tls}^{rel}	δx_{stln}^{rel}
e-10	6.1e-10	6.1e-10	1.9e-10	3.9e-11	3.9e-11	2.3e-11	8.4e-10	8.4e-10	3.5e-10
e-8	6.91e-8	6.91e-8	2.43e-8	4.65e-9	4.65e-9	2.56e-9	8.88e-8	8.88e-8	3.31e-8
e-6	7.31e-6	7.31e-6	1.97e-6	4.08e-7	4.08e-7	2.27e-7	7.00e-6	7.00e-6	2.94e-6
e-4	6.37e-4	6.37e-4	2.01e-4	3.78e-5	3.78e-5	2.45e-5	8.19e-4	8.15e-4	3.42e-4
5e-4	3.34e-3	3.30e-3	1.05e-3	2.02e-4	2.02e-4	1.21e-4	4.33e-3	4.20e-3	1.75e-3
e-3	6.38e-3	6.39e-3	2.02e-3	3.66e-4	3.58e-4	2.33e-4	8.26e-3	8.13e-3	3.32e-3
5e-3	4.53e-2	3.70e-2	1.06e-2	2.79e-3	2.10e-3	1.18e-3	7.57e-2	4.10e-2	1.71e-2
e-2	1.04e-1	6.67e-2	2.25e-2	7.12e-3	4.23e-3	2.22e-3	2.29e-1	8.02e-2	3.48e-2
e-1				3.99e-1	5.59e-2	2.70e-2			
	(a)			(b)			(c)		

converges to the optimal solution in model reduction problems for which the global optimum can be derived analytically.

The benefits of the STLN method in various parameter estimation problems, such as linear prediction, impulse response estimation of multivariable systems, and transfer function modeling, are studied. The L_2 norm STLN method is compared with the LS and TLS methods and shown to improve the accuracy of the parameter estimates by a factor 2 to 40 at any SNR.

In order to more fully investigate the potential of the STLN problem formulation and algorithms for a range of applications, a number of areas need further study. The benefits of the STLN approach and the computational performance of the STLN algorithms need to be investigated in much larger problems arising in actual applications, e.g., in signal enhancement and signal reconstruction and NMR data quantification. The STLN algorithms need further improvements in cases of near rank deficiency of A and ill conditioning of the kernel problem solved in each iteration. In addition, the convergence rate needs further analysis. More refined algorithms with better convergence properties are currently under study [15].

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. Signal Processing*, vol. 39, pp. 1070–1087, 1991.

[2] A. A. Anda and H. Park, "Self-scaling fast rotations for stiff least squares problems," *Linear Algebra, Applicat.*, vol. 234, pp. 137–161, 1996.
[3] H. Chen, S. Van Huffel, and J. Vandewalle, "Structured total least norm and its application to parameter estimation of exponentially damped sinusoids," in *Proc. ProRISC/IEEE Benelux Workshop Circuits, Syst. Signal Processing*, Mierlo, The Netherlands, Mar. 1995, pp. 45–50.
[4] J. Cadzow, "Signal enhancement: A composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 49–62, 1988.
[5] J. W. Demmel, "The smallest perturbation of a submatrix which lowers the rank and constrained total least squares problems," *SIAM J. Numer. Anal.*, vol. 24, pp. 199–206, 1987.
[6] B. De Moor, "Structured total least squares and L_2 approximation problems," *Linear Algebra, Applicat.*, Special Issue on *Numerical Linear Algebra Methods in Control, Signals and Systems*, Van Dooren *et al.*, Eds., vol. 188–189, pp. 163–207, 1993.
[7] B. L. R. De Moor, P. Van Overschee, and G. Schelfhout, "H_2 model reduction for SISO systems," in *Proc. IFAC '93 Conf.*, Sydney, Australia, vol. 2, 1993, pp. 227–230.
[8] B. De Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Trans. Signal Processing*, vol. 42, pp. 3104–3113, 1994.
[9] B. De Moor and B. Roorda, "L_2-optimal linear system identification—Structured total least squares for SISO systems," in *Proc. 33rd Conf. Decision Contr.*, Lake Buena Vista, Dec. 1994, pp. 2874–2879.
[10] R. Fletcher, *Practical Methods of Optimization, Vol. I, Unconstrained Optimization.* New York: Wiley, 1980.
[11] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their l_∞-error bounds," *Int. J. Contr*, vol. 39, pp. 1115–1193, 1984.
[12] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, pp. 883–893, 1980.
[13] ——, *Matrix Computations.* Baltimore, MD: Johns Hopkins University Press, 1989, 2nd ed.
[14] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1991, 2nd ed.
[15] P. Lemmerling and S. Van Huffel, "Superlinear algorithms for solving the structured total least norm problem," ESAT-SISTA Rep. TR 95-44, ESAT Lab., Katholieke Univ. Leuven, Belgium, Sept. 1995.
[16] K. Madsen, "An algorithm for minimax solution of overdetermined systems of nonlinear equations," *J. Inst. Math. Applications*, vol. 16, pp. 321–328, 1975.
[17] M. R. Osborne and G. A. Watson, "An analysis of the total approximation problem in separable norms and an algorithm for the total l_1 problem," *SIAM J. Sci. Stat. Comput.*, vol. 6, pp. 410–424, 1985.
[18] H. Park, J. B. Rosen, and J. Glick, "Structured total least norm method for Toeplitz problems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Detroit, MI, vol. 2, May 1995, pp. 1141–1144.
[19] M. A. Rahman and K. B. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1440–1454, 1987.
[20] J. B. Rosen, H. Park, and J. Glick, "The total least norm problem: Formulation and algorithms," Army High Performance Comput. Res. Center preprint 94-041, Univ. of Minnesota, Nov. 1993, revised July 1994.
[21] ——, "Total least norm formulation and solution for structured problems," *SIAM J. Matrix Anal. Appl.*, vol. 17, pp. 110–126, 1996.
[22] G. Schelfhout and B. De Moor, "Frequency weighted H_2 and Hilbert–Schmidt–Hankel norm model reduction," in *Proc. 33rd Conf. Decision Contr.*, Lake Buena Vista, Dec. 1994, pp. 3215–3216.
[23] G. Schelfhout, "Model reduction for control design," Ph.D. thesis, Dept. Elect. Eng., Katholieke Univ. Leuven, Belgium, Jan. 1996.
[24] H. Späth and G. A. Watson, "On orthogonal linear L_1 approximation," *Numer. Math.*, vol. 51, pp. 531–543, 1987.
[25] H. Späth, *Mathematical Algorithms for Linear Regression.* Boston: Academic, 1992.
[26] S. Van Huffel and J. Vandewalle, "Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns of A are subject to errors," *SIAM J. Matrix Anal. Appl.*, vol. 10, pp. 294–315, 1989.
[27] ——, *The Total Least Squares Problem: Computational Aspects and Analysis.* Philadelphia: SIAM, 1991.
[28] S. Van Huffel, B. De Moor, and H. Chen, "Relationships between constrained and structured TLS, with applications to signal enhancement," in *Systems and Networks: Mathematical Theory and Applications*, Mathematical Research, vol. 79, U. Helmke, R. Mennicken, and J. Saurer, Eds. Berlin: Akademie Verlag, 1994, vol. II, pp. 899–904.
[29] C. F. Van Loan, "On the method of weighting for equality constrained least squares problems," *SIAM J. Numer. Anal.*, vol. 22, pp. 851–864, 1985.
[30] G. A. Watson, "On a class of algorithms for total approximation," *J. Approx. Theory*, vol. 45, pp. 219–231, 1985.



Sabine Van Huffel (M'88) was born on September 26, 1958, received the M. Eng. degree in computer sciences in 1981, and received a post-graduate degree in biomedical engineering in 1985, both from the Katholieke Universiteit Leuven, Belgium. In April 1983, she moved to the ESAT Laboratory, Department of Electrical Engineering, where she received the degree of Doctor in Applied Sciences in June 1987.

At present, she is a research associate of the Belgian National Fund of Scientific Research (NFWO) and associate professor of the Katholieke Universiteit Leuven. Her research interests are in numerical linear algebra, errors-in-variables regression, system identification, parameter estimation, and biomedical signal processing. She is author of a book entitled *The Total Least Squares Problem: Computational Aspects and Analysis* (Philadelphia: SIAM, 1991).

Currently, Dr. Van Huffel is an associate editor of *Numerical Algorithms* and the *SIAM Journal on Matrix Analysis and Applications.*



Haesun Park (M'90) was born in Seoul, Korea. She received the B.S. degree in mathematics Summa Cum Laude from Seoul National University, Seoul, Korea, in 1981 and the M.S. and Ph.D. degrees in computer science from Cornell University, Ithaca, NY, in 1985 and 1987, respectively.

She has been an Associate Professor in the Department of Computer Science, University of Minnesota, Minneapolis, since 1993, where she was an Assistant Professor from 1987 to 1993. Her research interests include numerical linear algebra and parallel processing for scientific computing. More recently, she has investigated fast and accurate solution methods for recursive least squares problems and the solution of structured total least norm and parameter estimation problems. Her research is currently supported by the National Science Foundation.

Dr. Park is currently serving on the editorial board of the *SIAM Journal on Scientific Computing.*



J. Ben Rosen received the B.S. degree in electrical engineering from Johns Hopkins University, Baltimore, MD, in 1943 and the Ph.D. degree in applied mathematics from Columbia University, New York, NY, in 1952.

From 1953 to 1962, he was Head of the Applied Math Department at Shell Development Co., Emeryville, CA. He was Visiting Professor of Computer Science at Stanford University, Stanford, CA, from 1962 to 1964 and Professor of Computer Science at the University of Wisconsin, Madison, from 1964 to 1971. From 1971 to 1992, he was Professor at the Computer Science Department at the University of Minnesota, Minneapolis. He is now Professor Emeritus at the University of Minnesota, Adjunct Professor at the Computer Science and Engineering Department of the University of California at San Diego, and a Senior Fellow at both the San Diego Supercomputer Center and the Minnesota Supercomputer Institute. His research activity has been primarily in numerical optimization and related areas. More recently, he has investigated the solution of large-scale linear and nonlinear programming problems with special structure, parallel methods for numerical optimization, global optimization for predicting molecular structure, and computational methods for structured total least norm applications. He has supervised the research of 21 Ph.D. students, and his research is currently supported by the National Science Foundation.