STATISTICS AND ITS INTERFACE Volume ? (2010) 1–6

Bibliography on total least squares and related methods

IVAN MARKOVSKY

The class of total least squares methods has been growing since the basic total least squares method was proposed by Golub and Van Loan in the 70's. Efficient and robust computational algorithms were developed and properties of the resulting estimators were established in the errors-in-variables setting. At the same time the developed methods were applied in diverse areas, leading to broad literature on the subject. This paper collects the main references and guides the reader in finding details about the total least squares methods and their applications. In addition, the paper comments on similarities and differences between the total least squares and the singular spectrum analysis methods.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 15A06, 62J12; secondary 37M10.
KEYWORDS AND PHRASES: total least squares, errors-in-variables, singular spectrum analysis, applications.

1. INTRODUCTION

The term "total least squares" refers to a range of problems, solution methods, and algorithms aiming at an approximate solution of an overdetermined linear system of equations

$$(1) \qquad Ax \approx b, \qquad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m,\ m > n,$$

where both the right hand side b and the coefficients matrix A are perturbed. The basic total least squares approximate solution of (1) is defined by the following optimization problem:

$$(2) \qquad \begin{array}{ll} \text{minimize} & \text{over } \widehat{A}, \widehat{b}, \text{ and } x \quad \|A-\widehat{A}\|_F^2 + \|b-\widehat{b}\|_2^2 \\ \text{subject to} & \widehat{A}x = \widehat{b}, \end{array}$$

where $\|\cdot\|_F$ is the Frobenius norm and $\|\cdot\|_2$ is the Euclidean norm. The cost function measures the size of the perturbation, applied on the data, and the constraint ensures that the perturbed system has a solution. The solution of the optimally perturbed system is the total least squares approximate solution of the original incompatible system (1). Contrast (2) with the basic least squares problem

$$\begin{array}{ll} \text{minimize} & \text{over } \widehat{b} \text{ and } x \quad \|b-\widehat{b}\|_2^2 \\ \text{subject to} & Ax = \widehat{b}, \end{array}$$

where the coefficients matrix A is assumed exact and is therefore not perturbed.

In a statistical setting, total least squares problems correspond to errors-in-variables estimation problems:

$$(3) \qquad A = A_0 + \widetilde{A}, \quad b = b_0 + \widetilde{b}, \quad A_0x_0 = b_0.$$

Here x_0 is the true value of the parameter x; A_0 and b_0 are the true data values; and $\widetilde{A}$ and $\widetilde{b}$ are the measurement noises, which satisfy some additional assumptions. Contrast (3) with the classical regression problem—the statistical setup, corresponding to the least squares problem.

The main sources of information on the total least squares topic are

- *overview papers:* [104], [95], [54], [47], [49];
- *proceedings and special issues:* [94], [97], [96, 98]; and
- *books:* [102], [58].

In this paper, we give a bibliography of the main publications up to 2010 on total least squares and related methods. In Section 1 classic papers on the basic total least squares problem are cited. Sections 2 and 3 discuss the extensions of the basic problem to weighted and structured total least squares problems. Section 4 lists applications of the total least squares methods. Section 5 explains the similarities and differences between total least squares methods to the singular spectrum analysis methods.

2. THE BASIC TOTAL LEAST SQUARES PROBLEM

The basic total least squares problem (2) and its solution by the singular value decomposition was introduced by Golub and Van Loan in [25, 27]. Van Huffel and Vandewalle [100] considered multivariable and nongeneric cases, when the problem has no solution and generalized the algorithm of Golub and Van Loan to produce a solution in these cases. A novel theoretical and computational framework for treating non-generic and non-unique total least squares problems was presented by Paige and Strakos [69].

Statistical properties of the total least squares method were studied by Gleser [24], who proved that the method yields a consistent estimator for the true parameter value in the *errors-in-variables* setting [23, 9]. The noise assumptions that ensure consistency of the basic total least squares method imply that all elements of the data matrix are measured with equal precision, an assumption that may not be satisfied in practice.

A variation of the total least squares problem is the *data least squares* problem [16], where the A matrix is noisy and the b matrix is exact. When the errors are row-wise independent with equal row covariance matrix (which is known up to a scaling factor), the *generalised total least squares* problem formulation [101] extends the consistency of the basic total least squares estimator.

3. WEIGHTED TOTAL LEAST SQUARES PROBLEMS

In the basic total least squares problem (2), the perturbation size is measured by the Frobenius norm, which puts equal weight on all elements. Generalising the Frobenius norm to a weighted norm,

$$\|A-\widehat{A}\|_W := \sqrt{\operatorname{vec}^\top(A-\widehat{A})W\operatorname{vec}(A-\widehat{A})},$$

where W is a given $mn \times mn$ positive definite matrix and vec : $\mathbb{R}^{m\times n} \to \mathbb{R}^{mn}$ vectorizes a matrix, leads to what are called weighted total least squares problems. Special optimization methods for the weighted total least squares problem are proposed in [13, 107, 72, 48, 46, 53, 83]. The Riemannian singular value decomposition method of De Moor [13], has no proved convergence properties. The maximum likelihood principle component analysis method of Wentzell *et al.* [107] is an alternating least squares algorithm. It applies to the general weighted total least squares problem and is globally convergent, with linear convergence rate. The method of Premoli and Rastello [72] is a heuristic for solving the first order optimality condition. A solution of a nonlinear equation is sought instead of a minimum point of the original optimization problem. The method is locally convergent with super linear convergence rate. The region of convergence around a minimum point could be rather small in practice. The weighted low rank approximation framework of Manton *et al.* [46] proposes specialised optimization methods on a Grassman manifold. The least squares nature of the problem is not exploited by the algorithms proposed in [46]. Statistical properties of the weighted total least squares estimator are studied in [35].

4. STRUCTURED TOTAL LEAST SQUARES PROBLEMS

Another direction of generalizing the basic total least squares method is to impose structure constraints (*e.g.* Hankel, Toeplitz, and Sylvester structure) on the data matrix and its approximation:

$$\begin{aligned} &\text{minimize} \quad \text{over } \widehat{A}, \widehat{b}, \text{ and } x \quad \|A-\widehat{A}\|_F^2 + \|b-\widehat{b}\|_2^2 \\ &\text{subject to} \quad \widehat{A}x \approx \widehat{b} \quad \text{and} \quad \widehat{A} \text{ has the same structure as } A. \end{aligned}$$

Abatzoglou *et al.* [1] are considered to be the first who formulated a structured total least squares problem. They called their approach constrained total least squares and motivate the problem as an extension of the total least squares method to matrices with structure. The solution approach adopted by Abatzoglou *et al.* is closely related to the one of Aoki and Yue [2]. Again an equivalent optimization problem is derived, but it is solved numerically using a Newton-type optimization method.

Shortly after the publication of the work on the constrained total least squares problem, De Moor [13] lists many applications of the structured total least squares problem and outlines a new framework for deriving analytical properties and numerical methods. His approach is based on the Lagrange multipliers and the basic result is an equivalent problem, called the Riemannian singular value decomposition, which can be considered as a "nonlinear" extension of the classical singular value decomposition. As an outcome of the new problem formulation, an iterative solution method based on the inverse power iteration is proposed.

Another algorithm for solving the structured total least squares problem (even with ℓ_1 and ℓ_∞ norm in the cost function), called structured total least norm, is proposed by Rosen *et al.* [78]. In contrast to the approaches of Aoki, Yue and Abatzoglou *et al.*, Rosen *et al.* solve the problem in its original formulation, the constraint is linearised around the current iteration point, which results in a linearly constrained least squares problem. In the algorithm of Rosen *et al.*, the constraint is incorporated in the cost function by adding a multiple of its residual norm.

The weighted low rank approximation framework of Manton *et al.* [46] has been extended in [81, 82] to include structured low rank approximation problems. All problem formulations and solution methods cited above, except for the ones in the structured low rank approximation framework, aim at rank reduction of the data matrix by one. A generalization of the algorithm of Rosen *et al.* to problems with rank reduction by more than one is proposed by Van Huffel *et al.* [99]. It involves, however, Kronecker products that unnecessary inflate the dimension of the involved matrices.

When dealing with a general affine structure the constrained total least squares, Riemannian singular value decomposition, and structured total least norm methods have cubic computational complexity per iteration in the number of measurements. Fast algorithms with linear computational complexity are proposed by Mastronardi *et al.* [43, 62, 60], for special structured total least squares problems with data matrix $C = \begin{bmatrix} A & b \end{bmatrix}$ that is Hankel or composed of a Hankel block A and an unstructured column b. They use the structured total least norm approach but recognise that a matrix appearing in the kernel subproblem of the algorithm has low displacement rank. This structure is exploited using the Schur algorithm.

Efficient algorithms for problems with block-Hankel/Toeplitz structure and rank reduction with more than one are proposed by Markovsky *et al.* [56, 55, 52]. In addition, a numerically reliable and robust software implementation is available [50]. All methods, except for [99, 81, 82], reduce the rank of the data matrix by one. The efficiency varies from cubic for the methods that use a general affine structure to linear for the efficient methods of Lemmerling *et al.* [43] and Mastronardi *et al.* [62] that use a Hankel/Toeplitz type structure. Efficient solution methods for weighted structured total least squares problems are proposed in [52].

5. APPLICATIONS

Total least squares is applied in:

- system identification [77, 76, 42, 70, 59],
- linear system theory [15, 14],
- image reconstruction [73, 61, 21],
- speech and audio processing [44, 30],
- modal and spectral analysis [106, 108],
- chemometrics [107, 84],
- computer vision [66],
- machine learning [90],
- computer algebra [110, 51, 5], and
- astronomy [6].

Many problems in system identification and signal processing can be reduced to special types of block-Hankel and block-Toeplitz structured total least squares problems. An overview of errors-in-variables methods in system identification is given by Söderström in [89]. In [94, 97], the use of total least squares and errors-in-variables models in the application fields are surveyed and new algorithms that apply the total least squares concept are described.

In the field of signal processing, Cadzow [8], Bresler and Macovski [7] propose heuristic solution methods that turn out to be *suboptimal* with respect to the ℓ_2-optimality criterion, see Tufts and Shah [93] and De Moor [14, Section V]. These methods, however, became popular because of their simplicity. For example, the method of Cadzow is an iterative method that alternates between unstructured low rank approximation and structure enforcement, thereby only requiring singular value decomposition computations and manipulation of the matrix entries. For in-vivo magnetic resonance spectroscopy and audio coding, new state-space based methods have been derived by making use of the total least squares approach for spectral estimation with extensions to decimation and multichannel data quantification [39, 40].

Tufts and Shah propose in [93], a *noniterative* method for Hankel structured total least squares approximation that is based on perturbation analysis and provides a nearly optimal solution for high signal-to-noise ratio (SNR). In a statistical setting, this method achieves the Cramer-Rao lower bound asymptotically as the SNR tends to infinity. Noniterative methods for solving the linear prediction problem (which is equivalent to Hankel structured total least-squares problem) are proposed by Tufts and Kumaresan [92, 36].

Apart from systems, control and signal processing, total least squares applications emerge in other fields, including information retrieval [20], shape from moments [80], and computer algebra [110, 51].

6. SIMILARITY AND DIFFERENCES BETWEEN TOTAL LEAST SQUARES AND SINGULAR SPECTRUM ANALYSIS

Both singular spectrum analysis [29] and total least squares estimate a subspace that best fits the data (in the sense of minimising the approximation error's Frobenius norm). Indeed, the constraint $\widehat{A}x = \widehat{b}$ in the total least squares problem (2) implies that the approximating matrix $\begin{bmatrix}\widehat{A} & \widehat{b}\end{bmatrix} \in \mathbb{R}^{m\times(n+1)}$, $m > n$, has rank at most n. Therefore, the rows of $\begin{bmatrix}\widehat{A} & \widehat{b}\end{bmatrix}$ belong to a subspace of $\mathbb{R}^{n+1}$ with dimension at most n. (The multivariable version $AX \approx B$, $B \in \mathbb{R}^{m\times d}$, of the total least squares problem allows data fitting by a space of a general dimension r, $n \geq r \geq 1$.)

The core step in the singular spectrum analysis as well as in the basic total least squares method is the singular value decomposition of the data matrix ($\begin{bmatrix}A & b\end{bmatrix}$ is the total least squares case), followed by truncation of the smallest singular value(s). It is well known, see, *e.g.*, [91, Page 35, Theorem 5.8] that truncation of the d smallest singular values achieves optimal (in the Frobenius norm sense) rank reduction by d. Therefore, both methods have a common core subproblem: unstructured low-rank approximation of the data matrix.

Contrary to the singular spectrum analysis method, which is a nonparameteric technique, however, the ultimate goal of the total least squares method is parameter estimation (the parameter x in the linear model $Ax \approx b$). Therefore, after the estimation of the optimal fitting subspace, the total least squares algorithm involves an extra step of parameter computation (from a basis of the fitting subspace), which is not a part of the singular spectrum analysis. This extra step of parameter computation is the cause for the existence of nongeneric total least squares problems, *i.e.*, cases when the problem has no solution. In contrast, the singular spectrum analysis method always has a solution.

Singular spectrum analysis methods based on the linear recurrence formula [29, Chapter 5] are closely related to Hankel structured total least squares problems [54, Section 5]. The relation is due to the fact that time series satisfying a linear recurrence relation of order n, give rise to Hankel matrices of rank at most n. Therefore, Hankel structured low-rank approximation of a Hankel matrix constructed from a given time series leads to approximation with a time series satisfying a linear recurrence relation. Time series satisfying the linear recurrence formula are fundamental in signal processing, where many heuristic and local optimization methods for Hankel structured low-rank approximation are developed. Among the heuristics, a well known one is Cadzow's method [8].

In a summary, total least squares and singular spectrum analysis use the same technique (low-rank approximation of the data matrix) but aim at solutions to different problems. Total least squares solves a parameter estimation problem for a static linear input/output model $Ax = b$ and singular spectrum analysis aims at a broad range of problems, such as forecasting and smoothing of time series, detection of trends, periodicities, and structural changes.

ACKNOWLEDGEMENTS

Research supported by PinView (Personal Information Navigator adapting through VIEWing), an EU FP7 funded Collaborative Project 216529.

Received 3 January 2008

REFERENCES

[1] ABATZOGLOU, T., MENDEL, J. and HARADA, G. (1991). The constrained total least squares technique and its application to harmonic superresolution. *IEEE Trans. Signal Process.* **39** 1070–1087.

[2] AOKI, M. and YUE, P. (1970). On a priori error estimates of some identification methods. *IEEE Trans. Automat. Control* **15** 541–548.

[3] BECK, A. and BEN-TAL, A. (2006). On the Solution of the Tikhonov Regularization of the Total Least Squares. *SIAM J. Optimization* **17** 98–118.

[4] BECK, A., BEN-TAL, A. and TEBOULLE, M. (2006). Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares. *SIAM J. Matrix Anal. Appl.* **28** 425–445.

[5] BOTTING, B. (2004). Structured Total Least Squares for Approximate Polynomial Operations Master's thesis, School of Computer Science, University of Waterloo.

[6] BRANHAM, R. (1995). Multivariate orthogonal regression in astronomy. *Celestial mechanics & dynamical astronomy* **61** 239–251.

[7] BRESLER, Y. and MACOVSKI, A. (1986). Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Trans. Acust., Speech, Signal Process.* **34** 1081–1089.

[8] CADZOW, J. (1988). Signal enhancement—A composite property mapping algorithm. *IEEE Trans. Signal Process.* **36** 49–62.

[9] CARROLL, R., RUPPERT, D. and STEFANSKI, L. (1995). *Measurement Error in Nonlinear Models*. Chapman & Hall/CRC, London.

[10] CHEN, W., CHEN, M. and ZHOU, J. (2000). Adaptively regularized constrained total least-squares image restoration. *IEEE Trans. Image Process.* **9** 588–596.

[11] CHU, M., BUONO, N. D., LOPEZ, L. and POLITI, T. (2005). On the Low-Rank Approximation of Data on the Unit Sphere. *SIAM J. Matrix Anal. Appl.* **27** 46–60.

[12] CHU, M., FUNDERLIC, R. and PLEMMONS, R. (2003). Structured low rank approximation. *Linear Algebra Appl.* **366** 157–172.

[13] DE MOOR, B. (1993). Structured total least squares and L_2 approximation problems. *Linear Algebra Appl.* **188–189** 163–207.

[14] DE MOOR, B. (1994). Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Trans. Signal Process.* **42** 3104–3113.

[15] DE MOOR, B. and DAVID, J. (1992). Total linear least squares and the algebraic Riccati equation. *Control Lett.* **18** 329–337.

[16] DEGROAT, R. and DOWLING, E. (1991). The data least squares problem and channel equalization. *IEEE Trans. Signal Process.* **41** 407–411.

[17] ECKART, G. and YOUNG, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika* **1** 211–218.

[18] FAZEL, M. (2002). Matrix Rank Minimization with Applications PhD thesis, Elec. Eng. Dept, Stanford University.

[19] FIERRO, R., GOLUB, G., HANSEN, P. and O'LEARY, D. (1997). Regularization by Truncated total least squares. *SIAM J. Sci. Comp.* **18** 1223–1241.

[20] FIERRO, R. and JIANG, E. (2005). Lanczos and the Riemannian SVD in Information Retrieval Applications. *Numer. Linear. Algebra Appl.* **12** 355–372.

[21] FU, H. and BARLOW, J. (2004). A regularized structured total least squares algorithm for high-resolution image reconstruction. *Linear Algebra Appl.* **391** 75–98.

[22] FU, H., NG, M. K. and BARLOW, J. L. (2006). Structured Total Least Squares for Color Image Restoration. *SIAM J. Sci. Comput.* **28** 1100–1119.

[23] FULLER, W. (1987). *Measurement Error Models*. Wiley, New York.

[24] GLESER, L. (1981). Estimation in a multivariate "Errors in variables" regression model: large sample results. *The Annals of Statistics* **9** 24–44.

[25] GOLUB, G. (1973). Some modified matrix eigenvalue problems. *SIAM Review* **15** 318–344.

[26] GOLUB, G., HANSEN, P. and O'LEARY, D. (1999). Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl.* **21** 185–194.

[27] GOLUB, G. and VAN LOAN, C. (1980). An analysis of the total least squares problem. *SIAM J. Numer. Anal.* **17** 883–893.

[28] GOLUB, G. H., HANSEN, P. C. and O'LEARY, D. P. (1999). Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl.* **21** 185–194.

[29] GOLYANDINA, N., NEKRUTKIN, V. and ZHIGLJAVSKY, A. (2001). *Analysis of time series structure*. CRC Press.

[30] HERMUS, K., VERHELST, W., LEMMERLING, P., WAMBACQ, P. and VAN HUFFEL, S. (2005). Perceptual audio modeling with exponentially damped sinusoids. *Signal Processing* **85** 163–176.

[31] HO, B. L. and KALMAN, R. E. (1966). Effective construction of linear state-variable models from input/output functions. *Regelungstechnik* **14** 545–592.

[32] KALSI, A. and O'LEARY, D. P. (2006). Algorithms for Structured Total Least Squares Problems with Applications to Blind Image Deblurring. *J. Res. Natl. Inst. Stan.* **111** 113–119.

[33] KAMM, J. and NAGY, J. G. (1997). Least Squares and Total Least Squares Methods in Image Restoration. In *WNAA '96: Proc. 1st Intl. Workshop Num. Anal. and Its Appl.* 212–219. Springer-Verlag, London, UK.

[34] KUKUSH, A., MARKOVSKY, I. and VAN HUFFEL, S. (2005). Consistency of the structured total least squares estimator in a multivariate errors-in-variables model. *J. Statist. Plann. Inference* **133** 315–358.

[35] KUKUSH, A. and VAN HUFFEL, S. (2004). Consistency of elementwise-weighted total least squares estimator in a multivariate errors-in-variables model $AX = B$. *Metrika* **59** 75–97.

[36] KUMARESAN, R. and TUFTS, D. (1982). Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Trans. Acust., Speech, Signal Process.* **30** 833–840.

[37] KUNG, S. (1978). A new identification method and model reduction algorithm via singular value decomposition. In *Proceedings of the 12th Asilomar Conference on Circuits, Systems, and Computers* 705–714.

[38] LAMPE, J. and VOSS, H. (2007). On a quadratic eigenproblem occurring in regularized total least squares. *Comput. Stat. Data Anal.* **52** 1090–1102.

[39] LAUDADIO, T., MASTRONARDI, N., VANHAMME, L., VAN HECKE, P. and VAN HUFFEL, S. (2002). Improved Lanczos algorithms for blackbox MRS data quantitation. *Journal of Magnetic Resonance* **157** 292–297.

[40] LAUDADIO, T., SELEN, Y., VANHAMME, L., STOICA, P., VAN HECKE, P. and VAN HUFFEL, S. (2004). Subspace-based MRS data quantitation of multiplets using prior knowledge. *Journal of Magnetic Resonance* **168** 53–65.

[41] LEI, J., LIU, S., LI, Z., SCHLABERG, H. and SUN, M. (2008). An image reconstruction algorithm based on the regularized total least squares method for electrical capacitance tomography. *Flow Meas. Instrum.* **19** 325–330.

[42] LEMMERLING, P. and DE MOOR, B. (2001). Misfit versus latency. *Automatica* **37** 2057–2067.

[43] LEMMERLING, P., MASTRONARDI, N. and VAN HUFFEL, S. (2000). Fast algorithm for solving the Hankel/Toeplitz structured total least squares problem. *Numerical Algorithms* **23** 371–392.

[44] LEMMERLING, P., MASTRONARDI, N. and VAN HUFFEL, S. (2003). Efficient implementation of a structured total least squares based speech compression method. *Linear Algebra Appl.* **366** 295–315.

[45] LEMMERLING, P., VAN HUFFEL, S. and DE MOOR, B. (2002). The structured total least squares approach for nonlinearly structured matrices. *Numer. Linear. Algebra Appl.* **9** 321–332.

[46] MANTON, J., MAHONY, R. and HUA, Y. (2003). The geometry of weighted low-rank approximations. *IEEE Trans. Signal Process.* **51** 500–514.

[47] MARKOVSKY, I. (2007). Structured low-rank approximation and its applications. *Automatica* **44** 891–909.

[48] MARKOVSKY, I., RASTELLO, M.-L., PREMOLI, A., KUKUSH, A. and VAN HUFFEL, S. (2005). The element-wise weighted total least squares problem. *Comput. Statist. Data Anal.* **50** 181–209.

[49] MARKOVSKY, I., SIMA, D. M. and VAN HUFFEL, S. (2010). Total least squares methods. *Wiley Interdisciplinary Reviews: Computational Statistics* **2** 212–217.

[50] MARKOVSKY, I. and VAN HUFFEL, S. (2005). High-performance numerical algorithms and software for structured total least squares. *J. of Comput. and Appl. Math.* **180** 311–331.

[51] MARKOVSKY, I. and VAN HUFFEL, S. (2006). An algorithm for approximate common divisor computation. In *Proceedings of the 17th Symposium on Mathematical Theory of Networks and Systems* 274–279.

[52] MARKOVSKY, I. and VAN HUFFEL, S. (2006). On weighted structured total least squares. In *Large-Scale Scientific Computing* (I. LIRKOV, S. MARGENOV and J. WAŚNIEWSKI, eds.). *Lecture Notes in Computer Science* **3743** 695–702. Springer–Verlag, Berlin.

[53] MARKOVSKY, I. and VAN HUFFEL, S. (2007). Left vs right representations for solving weighted low rank approximation problems. *Linear Algebra Appl.*

422 540–552.
[54] MARKOVSKY, I. and VAN HUFFEL, S. (2007). Overview of total least squares methods. *Signal Processing* **87** 2283–2302.
[55] MARKOVSKY, I., VAN HUFFEL, S. and KUKUSH, A. (2004). On the computation of the structured total least squares estimator. *Numer. Linear. Algebra Appl.* **11** 591–608.
[56] MARKOVSKY, I., VAN HUFFEL, S. and PINTELON, R. (2005). Block-Toeplitz/Hankel structured total least squares. *SIAM J. Matrix Anal. Appl.* **26** 1083–1099.
[57] MARKOVSKY, I., WILLEMS, J. C. and DE MOOR, B. (2006). Comparison of identification algorithms on the database for system identification DAISY. In *Proceedings of the 17th Symposium on Mathematical Theory of Networks and Systems* 2858–2869.
[58] MARKOVSKY, I., WILLEMS, J. C., VAN HUFFEL, S. and DE MOOR, B. (2006). *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. *Monographs on Mathematical Modeling and Computation* 11. SIAM.
[59] MARKOVSKY, I., WILLEMS, J. C., VAN HUFFEL, S., MOOR, B. D. and PINTELON, R. (2005). Application of structured total least squares for system identification and model reduction. *IEEE Trans. Automat. Control* **50** 1490–1500.
[60] MASTRONARDI, N. (2001). Fast and reliable algorithms for structured total least squares and related matrix problems PhD thesis, ESAT/SISTA, K.U. Leuven.
[61] MASTRONARDI, N., LEMMERLING, P., KALSI, A., O'LEARY, D. and VAN HUFFEL, S. (2004). Implementation of the regularized structured total least squares algorithms for blind image deblurring. *Linear Algebra Appl.* **391** 203–221.
[62] MASTRONARDI, N., LEMMERLING, P. and VAN HUFFEL, S. (2000). Fast structured total least squares algorithm for solving the basic deconvolution problem. *SIAM J. Matrix Anal. Appl.* **22** 533–553.
[63] MASTRONARDI, N., LEMMERLING, P. and VAN HUFFEL, S. (2005). Fast regularized structured total least squares algorithm for solving the basic deconvolution problem. *Numer. Linear. Algebra Appl.* **12** 201–209.
[64] MASTRONARDI, N., LEMMERLING, P. and VAN HUFFEL, S. (2005). Fast regularized structured Total Least Squares problems for solving the basic deconvolution problem. *Numer. Linear. Algebra Appl.* **12** 201–209.
[65] MESAROVIĆ, V., GALATSANOS, N. and KATSAGGELOS, A. (1995). Regularized constrained total least squares image restoration. *IEEE Trans. Image Process.* **4** 1096–1108.
[66] MÜHLICH, M. and MESTER, R. (1998). The role of total least squares in motion analysis. In *Proceedings of the 5th European Conference on Computer Vision* (H. BURKHARDT, ed.) 305–321. Springer-Verlag.
[67] NG, M., KOO, J. and BOSE, N. (2002). Constrained total least squares computations for high resolution image reconstruction with multisensors. *Int. J. Imaging Systems Technol.* **12** 35–42.
[68] NG, M., PLEMMONS, R. and PIMENTEL, F. (2000). A new approach to constrained total least squares image restoration. *Linear Algebra Appl.* **316** 237–258.
[69] PAIGE, C. and STRAKOS, Z. (2005). Core problems in linear algebraic systems. *SIAM J. Matrix Anal. Appl.* **27** 861–875.
[70] PINTELON, R., GUILLAUME, P., VANDERSTEEN, G. and ROLAIN, Y. (1998). Analyses, development, and applications of TLS algorithms in frequency domain system identification. *SIAM J. Matrix Anal. Appl.* **19** 983–1004.
[71] PINTELON, R. and SCHOUKENS, J. (2001). *System Identification: A Frequency Domain Approach*. IEEE Press, Piscataway, NJ.
[72] PREMOLI, A. and RASTELLO, M.-L. (2002). The parametric quadratic form method for solving TLS problems with elementwise weighting. In *Total least squares and errors-in-variables modeling: Analysis, algorithms and applications* (S. VAN HUFFEL and P. LEMMERLING, eds.) 67–76. Kluwer.
[73] PRUESSNER, A. and O'LEARY, D. (2003). Blind deconvolution using a regularized structured total least norm algorithm. *SIAM J. Matrix Anal. Appl.* **24** 1018–1037.
[74] PRUESSNER, A. and O'LEARY, D. P. (2003). Blind Deconvolution Using a Regularized Structured Total Least Norm Approach. *SIAM J. Matrix Anal. Appl.* **24** 1018–1037.
[75] RENAUT, R. and GUO, H. (2005). Efficient algorithms for solution of regularized total least squares. *SIAM J. Matrix Anal. Appl.* **26** 457–476.
[76] ROORDA, B. (1995). Algorithms for global total least squares modelling of finite multivariable time series. *Automatica* **31** 391–404.
[77] ROORDA, B. and HEIJ, C. (1995). Global total least squares modeling of multivariate time series. *IEEE Trans. Automat. Control* **40** 50–63.
[78] ROSEN, J., PARK, H. and GLICK, J. (1996). Total least norm formulation and solution of structured problems. *SIAM J. Matrix Anal. Appl.* **17** 110–126.
[79] ROSEN, J., PARK, H. and GLICK, J. (1998). Structured total least norm for nonlinear problems. *SIAM J. Matrix Anal. Appl.* **20** 14–30.
[80] SCHUERMANS, M., LEMMERLING, P., LATHAUWER, L. D. and VAN HUFFEL, S. (2006). The use of total least squares data fitting in the shape from moments problem. *Signal Processing* **86** 1109–1115.
[81] SCHUERMANS, M., LEMMERLING, P. and VAN HUFFEL, S. (2004). Structured weighted low rank approximation. *Numer. Linear. Algebra Appl.* **11** 609–618.
[82] SCHUERMANS, M., LEMMERLING, P. and VAN HUFFEL, S. (2006). Block-row Hankel Weighted Low Rank Approximation. *Numer. Linear. Algebra Appl.* **13** 293–302.
[83] SCHUERMANS, M., MARKOVSKY, I. and VAN HUFFEL, S. (2007). An adapted version of the element-wise weighted TLS method for applications in chemometrics. *Chemometrics and Intelligent Laboratory Systems* **85** 40–46.
[84] SCHUERMANS, M., MARKOVSKY, I., WENTZELL, P. and VAN HUFFEL, S. (2005). On the equivalence between total least squares and maximum likelihood PCA. *Analytica Chimica Acta* **544** 254–267.
[85] SIMA, D. and VAN HUFFEL, S. (2004). Appropriate cross-validation for regularized errors-in-variables linear models. In *Proceedings of the COMPSTAT 2004 Symposium*. Physica–Verlag/Springer, Prague, Czech Republic.
[86] SIMA, D., VAN HUFFEL, S. and GOLUB, G. (2004). Regularized Total Least Squares based on quadratic eigenvalue problem solvers. *BIT Numerical Mathematics* **44** 793–812.
[87] SIMA, D. M. and VAN HUFFEL, S. (2007). Level choice in truncated total least squares. *Comput. Stat. Data Anal.* **52** 1103–1118.
[88] SIMA, D. M., VAN HUFFEL, S. and GOLUB, G. H. (2004). Regularized total least squares based on quadratic eigenvalue problem solvers. *BIT* **44** 793–812.
[89] SÖDERSTRÖM, T. (2006). Errors-in-variables methods in system identification. In *14th IFAC Symposium on System Identification* 29–41.
[90] SREBRO, N. (2004). Learning with Matrix Factorizations PhD thesis, MIT.
[91] TREFETHEN, L. and BAU, D. (1997). *Numerical Linear Algebra*. SIAM.
[92] TUFTS, D. and KUMARESAN, R. (1982). Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood. *Proc. of the IEEE* **70** 975–989.
[93] TUFTS, D. and SHAH, A. (1993). Estimation of a signal waveform from noisy data using low-rank approximation to a data matrix. *IEEE Trans. Signal Process.* **41** 1716–1721.
[94] VAN HUFFEL, S., ed. (1997). *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*. SIAM, Philadelphia.
[95] VAN HUFFEL, S. (2004). Total least squares and Errors-in-Variables Modeling: Bridging the gap between Statistics, Computational Mathematics and Engineering. In *COMPSTAT Proceedings in Computational Statistics* (J. Antoch, ed.) 539–555. Physika-Verlag, Heidelberg.
[96] VAN HUFFEL, S., CHENG, C.-L., MASTRONARDI, N., PAIGE, C. and KUKUSH, A. (2007). Editorial: Total least squares and errors-in-variables modeling. *Comput. Statist. Data Anal.* **52** 1076–1079.
[97] VAN HUFFEL, S. and LEMMERLING, P., eds. (2002). *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*. Kluwer.
[98] VAN HUFFEL, S., MARKOVSKY, I., VACCARO, R. J. and SÖDERSTRÖM, T. (2007). Guest editorial: Total least squares and errors-in-variables modeling. *Signal Processing* **87** 2281–2282.
[99] VAN HUFFEL, S., PARK, H. and ROSEN, J. (1996). Formulation and solution of structured total least norm problems for parameter estimation. *IEEE Trans. Signal Process.* **44** 2464–2474.
[100] VAN HUFFEL, S. and VANDEWALLE, J. (1988). Analysis and solution of the nongeneric total least squares problem. *SIAM J. Matrix Anal. Appl.* **9** 360–372.
[101] VAN HUFFEL, S. and VANDEWALLE, J. (1989). Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. *SIAM J. Matrix Anal.* **10** 294–315.

[102] VAN HUFFEL, S. and VANDEWALLE, J. (1991). *The total least squares problem: Computational aspects and analysis*. SIAM, Philadelphia.

[103] VAN HUFFEL, S. and ZHA, H. (1991). The restricted total least squares problem: Formulation, algorithm and properties. *SIAM J. Matrix Anal. Appl.* **12** 292–309.

[104] VAN HUFFEL, S. and ZHA, H. (1993). The total least squares problem. In *Handbook of Statistics: Computational Statistics*, (C. R. Rao, ed.) **9** 377–408. Elsevier Science Publishers B.V., Amsterdam.

[105] VANLUYTEN, B., WILLEMS, J. C. and DE MOOR, B. (2005). Model Reduction of Systems with Symmetries. In *Proc. of the 44th IEEE Conference on Decision and Control* 826–831.

[106] VERBOVEN, P., GUILLAUME, P., CAUBERGHE, B., PARLOO, E. and VANLANDUIT, S. (2004). Frequency-domain generalized total least-squares identification for modal analysis. *Journal of Sound and Vibration* **278** 21–38.

[107] WENTZELL, P., ANDREWS, D., HAMILTON, D., FABER, K. and KOWALSKI, B. (1997). Maximum likelihood principal component analysis. *J. Chemometrics* **11** 339–366.

[108] YEREDOR, A. (2004). Multiple delays estimation for chirp signals using structured total least squares. *Linear Algebra Appl.* **391** 261–286.

[109] YOUNAN, N. and FAN, X. (1998). Signal restoration via the regularized constrained total least squares. *Signal Processing* **71** 85–93.

[110] ZHI, L. and YANG, Z. (2004). Computing Approximate GCD of Univariate Polynomials by Structure Total Least Norm. In *MM Research Preprints* 24 375–387. Academia Sinica.

Ivan Markovsky
School of Electronics and Computer Science
University of Southampton
Southampton, SO17 1BJ, UK
E-mail address: im@ecs.soton.ac.uk