Article type: Advanced Review

Generalizations of the total least squares problem **Article ID**

Ivan Markovsky

School of Electronics and Computer Science, University of Southampton, U.K.

Diana M. Sima, Sabine Van Huffel

Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium

Keywords

total least squares, errors-in-variables, data modeling, regularization

Abstract

Recent advances in total least squares approaches for solving various errors-in-variables modeling problems are reviewed, with emphasis on the following generalizations:

1. the use of weighted norms as a measure of the data perturbation size, capturing prior knowledge about uncertainty in the data;
2. the addition of constraints on the perturbation to preserve the structure of the data matrix, motivated by structured data matrices occurring in signal and image processing, systems and control, and computer algebra;
3. the use of regularization in the problem formulation, aiming at stabilizing the solution by decreasing the effect due to intrinsic ill-conditioning of certain problems.

For extensive reviews of the total least squares approach and its applications, we refer the reader to the following

- *overview papers:* [Van Huffel and Zha, 1993], [Van Huffel, 2004], [Markovsky and Van Huffel, 2007a], [Markovsky, 2007];
- *proceedings and special issues:* [Van Huffel, 1997], [Van Huffel and Lemmerling, 2002], [Van Huffel et al., 2007a,b]; and
- *books:* [Van Huffel and Vandewalle, 1991], [Markovsky et al., 2006].

The focus of this paper is on computational algorithms for solving the generalized total least squares problems. The reader is referred to the errors-in-variables literature for the statistical properties of the corresponding estimators, as well as for a wider range of applications.

Weighted and structured total least squares problems

The TLS solution

$$\widehat{x}_{\text{tls}} = \arg\min_{x,\widehat{A},\widehat{b}} \left\| \begin{bmatrix} A & b \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{b} \end{bmatrix} \right\|_{\text{F}} \quad \text{subject to} \quad \widehat{A}x = \widehat{b} \tag{1}$$

of an overdetermined system of equations $Ax \approx b$ is appropriate when all elements of the data matrix $\begin{bmatrix} A & b \end{bmatrix}$ are noisy and the noise is zero mean, independent and identically distributed. More precisely, (under regularity conditions) $\widehat{x}_{\text{tls}}$ is a consistent estimator for the true parameter value $\overline{x}$ in the *errors-in-variables (EIV) model*

$$A = \overline{A} + \widetilde{A}, \qquad b = \overline{b} + \widetilde{b}, \qquad \overline{A}\overline{x} = \overline{b}, \tag{2}$$

where the vector of perturbations $\operatorname{vec}(\begin{bmatrix} \widetilde{A} & \widetilde{b} \end{bmatrix})$ is zero mean and has covariance matrix that is equal to the identity up to a scaling factor, *i.e.*,

$$\mathbf{E}\left(\operatorname{vec}(\begin{bmatrix} \widetilde{A} & \widetilde{b} \end{bmatrix})\right) = 0 \qquad \text{and} \qquad \operatorname{cov}\left(\operatorname{vec}(\begin{bmatrix} \widetilde{A} & \widetilde{b} \end{bmatrix})\right) = \sigma^2 I. \tag{3}$$

The noise assumption (3) implies that all elements of the data matrix are measured with equal precision, an assumption that may not be satisfied in practice.

A natural generalization of the EIV model (2,3) is to allow the covariance matrix of the vectorized noise to be of the form $\sigma^2 V$, where V is a given positive definite matrix. The corresponding estimation problem is the TLS problem (1) with the Frobenius norm $\|\cdot\|_{\text{F}}$ replaced by the weighted matrix norm

$$\|\Delta D\|_{V^{-1}} := \sqrt{\operatorname{vec}^\top(\Delta D) V^{-1} \operatorname{vec}(\Delta D)}$$

i.e.,

$$\min_{x,\widehat{A},\widehat{b}} \left\| \begin{bmatrix} A & b \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{b} \end{bmatrix} \right\|_{V^{-1}} \quad \text{subject to} \quad \widehat{A}x = \widehat{b}. \tag{4}$$

In [De Moor, 1993, Section 4.3] this problem is called *weighted total least squares (WTLS)*. Closely related to the WTLS problem are the weighted low rank approximation problem [Manton et al., 2003, Markovsky and Van Huffel, 2007b] and the maximum likelihood principal component analysis problem [Wentzell et al., 1997, Schuermans et al., 2005].

As opposed to the weighted least squares problem, which is a trivial generalization of classical least squares, the WTLS problem does not have, in general, a closed form solution similar to the one of the TLS problem. The most general WTLS problem with analytic solution has a weight matrix of the form $V^{-1} = V_2^{-1} \otimes V_1^{-1}$, where $\otimes$ is the Kronecker product, V_1 is $m \times m$, and V_2 is $(n+1) \times (n+1)$ (m is the number of equations and n is the number of unknowns in $Ax \approx b$). For general weight matrix, the problem can be solved by local optimization methods. However, there is no guarantee that a globally optimal solution will be found.

There are two main categories of local optimization methods for solving WTLS problems: alternating projections and variable projections [Golub and Pereyra, 2003]. They are based on the observation that the constraint of the WTLS problem (4) is bilinear, which implies that the problem is linear in either x or $\widehat{A}$ and, therefore, can be solved globally and efficiently. Alternating projections is an iterative optimization algorithm that on each iteration step

1. solves a (linear) least squares problem in an $n \times (n+1)$ extended parameter X_{ext} with $\widehat{A}$ fixed to the value obtained on the previous iteration step

$$\min_{X_{\text{ext}}} \left\| \begin{bmatrix} A & b \end{bmatrix} - \widehat{A} X_{\text{ext}} \right\|_{V^{-1}}, \tag{5}$$

2. solves a least squares problem in $\widehat{A}$ with X_{ext} fixed to the optimal value of (5)

$$\min_{\widehat{A}} \left\| \begin{bmatrix} A & b \end{bmatrix} - \widehat{A} X_{\text{ext}} \right\|_{V^{-1}}. \tag{6}$$

The parameter x is recovered from X_{ext}, as follows

$$x := X_{\text{ext},1}^{-1} x_{\text{ext},2}, \qquad \text{where} \quad X_{\text{ext}} = \begin{bmatrix} \overbrace{X_{\text{ext},1}}^{n} & \overbrace{x_{\text{ext},2}}^{1} \end{bmatrix}.$$

In the statistical literature, the alternating projections algorithm is given the interpretation of *expectation maximization (EM)*. The problem of computing the optimal approximation $\widehat{A}$ given X_{ext} is the expectation step and the problem of computing X_{ext}, given $\widehat{A}$ is the maximization step of the EM procedure.

The variable projections method uses the closed form solution of the expectation problem (6):

$$f(x) := \sqrt{d^\top \left(V^{-1} - V^{-1} \mathbf{X}_{\text{ext}}{}^\top (\mathbf{X}_{\text{ext}} V^{-1} \mathbf{X}_{\text{ext}}{}^\top)^{-1} \mathbf{X}_{\text{ext}} V^{-1} \right) d},$$

where

$$d := \operatorname{vec}(\begin{bmatrix} A & b \end{bmatrix}) \qquad \text{and} \qquad \mathbf{X}_{\text{ext}} := \begin{bmatrix} I_n & x \end{bmatrix} \otimes I_m.$$

This is a projection of the rows of $\begin{bmatrix} A & b \end{bmatrix}$ on the subspace perpendicular to $\begin{bmatrix} x \\ -1 \end{bmatrix}$. The minimization over x is then an unconstrained nonlinear least squares problem $\min_x f(x)$, which can be solved by standard optimization methods, *e.g.*, the Levenberg-Marquardt method.

Another generalization of the TLS problem (1) is to add constraints on the approximation matrix $\begin{bmatrix} \widehat{A} & \widehat{b} \end{bmatrix}$ [Abatzoglou et al., 1991]. Such constraints are needed in applications where the data matrix is structured and the approximation is required to have the same structure. For example, in signal processing the output y of a finite impulse response (FIR) system to an input u is given by multiplication of a Toeplitz matrix constructed from u by the vector of the impulse response samples. In an FIR system estimation problem, where both the input and the output are noisy, the approximation

matrix is required to have Toeplitz structure in order for the result to have interpretation as a description of an FIR system.

Similarly to the WTLS problems, in general, *structured total least squares (STLS)* problems [De Moor, 1993] have no analytic solution in terms of the SVD of the data matrix. Beck and Ben-Tal [2006a] solved globally STLS problems with block-circulant structure by using the discrete Fourier transform and the solution of standard TLS problems. For other types of structure one has to resort to local optimization methods. In case of linearly structured problems, the constraint of the STLS optimization problem is bilinear, so that the alternating projections and variable projections methods, similar to the ones developed for the WTLS problem, can be used.

Regularized and truncated total least squares problems

Linear approximation problems $Ax \approx b$ are considered ill-posed when small variations in the data A and b lead to large variations in the computed solution x. In the context of ordinary least squares, methods such as ridge regression, Tikhonov regularization [Tikhonov and Arsenin, 1977] or truncated SVD [Hansen, 1990] are often employed to stabilize the computations. In recent years, several regularization formulations have also been explored in the context of the total least squares problem. We distinguish between methods based on penalties/constraints, and methods based on truncation.

The basic idea of *regularized total least squares (RTLS)* is forcing an upper bound on a weighted 2-norm of the solution vector x (although other types of constraints can be envisaged). Several formulations have been considered. A first formulation is the quadratically constrained RTLS problem stated in Golub et al. [1999], Sima et al. [2004], Renaut and Guo [2005], Beck et al. [2006] as

$$\min_{x,\widehat{A},\widehat{b}} \| \begin{bmatrix} A & b \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{b} \end{bmatrix} \|_F^2 \qquad \text{subject to } \widehat{A}x = \widehat{b},\ \|Lx\|_2^2 \leq \delta^2, \tag{7}$$

or, equivalently,

$$\min_{x} \frac{\|Ax-b\|_2^2}{1+\|x\|_2^2} \qquad \text{subject to } \|Lx\|_2^2 \leq \delta^2, \tag{8}$$

where L is a p by n matrix, usually the identity matrix or a discrete difference operator, and δ is a given scalar value.

A second formulation adds a Tikhonov-like quadratic penalty term $\|Lx\|_2^2$ to the TLS objective function [Beck and Ben-Tal, 2006b]:

$$\min_{x} \frac{\|Ax-b\|_2^2}{1+\|x\|_2^2} + \lambda \|Lx\|_2^2. \tag{9}$$

For δ^2 small enough (*i.e.*, $\delta^2 < \|L\widehat{x}_{\text{tls}}\|_2^2$ where $\widehat{x}_{\text{tls}}$ is the TLS solution), there exists a value of the parameter $\lambda > 0$ such that the solution of (8) coincides with the

solution of (9). A sufficient condition for attainability of the minima in (8) or (9) is: $\sigma_{\min}\left(\begin{bmatrix} AN & b \end{bmatrix}\right) < \sigma_{\min}(AN)$, where the columns of N form a basis for the nullspace of $L^\top L$ [Sima et al., 2004, Beck et al., 2006, Beck and Ben-Tal, 2006b].

As opposed to classical regularization methods in the context of ordinary least squares, these formulations do not have closed-form solutions. Although local optimization methods are used in practice, the analysis in Beck et al. [2006], Beck and Ben-Tal [2006b] suggests that both formulations can be recast in a global optimization framework, namely into scalar minimization problems, where each function evaluation requires the solution of a quadratically constrained linear least squares problem [Gander, 1981].

The constrained formulation (8) has been solved via a sequence of quadratic eigenvalue problems by Sima et al. [2004]. Combining this approach with the nonlinear Arnoldi method and reusing information from all previous quadratic eigenvalue problems, a more efficient method for large RTLS problems has been proposed in Lampe and Voss [2007]. Further, Renaut and Guo [2005] suggested an iterative method based on a sequence of linear eigenvalue problems, which has also been accelerated by solving the linear eigenproblems by the nonlinear Arnoldi method and by a modified root finding method based on rational interpolation [Lampe and Voss, 2008].

For the quadratic penalty formulation (9), a complete analysis has been presented by Beck and Ben-Tal [2006b]. A simple reformulation into a scalar minimization makes the problem more tractable:

$$\min_{\alpha} \ \mathcal{G}(\alpha), \ \text{where } \mathcal{G}(\alpha) := \min_{\|x\|_2^2=\alpha-1} \left\{ \frac{\|Ax-b\|_2^2}{\alpha} + \lambda \|Lx\|_2^2 \right\}. \tag{10}$$

In Lu et al. [2008] another related formulation called dual RTLS is proposed. It minimizes the norm $\|Lx\|_2^2$ subject to compatibility of the corrected system, as well as to upper bounds on $\|A-\widehat{A}\|_{\mathrm{F}}$ and $\|b-\widehat{b}\|_2$.

Truncation methods are another class of methods for regularizing linear ill-posed problems in the presence of measurement errors. In essence, they aim at limiting the contribution of noise or rounding errors by cutting off a certain number of terms in an SVD expansion. The *truncated total least squares* (TTLS) solution with truncation level k is the minimum 2-norm solution of $A_k x = b_k$, where $\begin{bmatrix} A_k & b_k \end{bmatrix}$ is the best rank-k approximation of $\begin{bmatrix} A & b \end{bmatrix}$. More precisely, if $U\Sigma V^\top$ is the SVD of $\begin{bmatrix} A & b \end{bmatrix}$,

$$x_{\mathrm{TTLS},k} = -V_{12}^k (V_{22}^k)^\dagger = -V_{12}^k (V_{22}^k)^\top / \|V_{22}^k\|^2, \tag{11}$$

where we partition V as (with $l = n-k+1$):

$$V = \begin{array}{c} \overset{\longleftarrow k \longrightarrow}{\overset{\longleftarrow l \longrightarrow}{\begin{bmatrix} V_{11}^k & V_{12}^k \\ V_{21}^k & V_{22}^k \end{bmatrix}}} \\ \end{array} \begin{array}{c} \updownarrow \ n \\ \updownarrow \ 1 \end{array} \tag{12}$$

The regularizing properties of truncated total least squares and a filter factor expansion of the TTLS solution have been described by Fierro et al. [1997]. Sima and Van Huffel

[2007] showed that the filter factors associated with the TTLS solution provide more information for choosing the truncation level compared to truncated SVD, where the filter factors are simply zeros and ones.

Applications and current trends

Core problem. The concept of core problem in linear algebraic systems has been developed by Paige and Strakoš [2005]. The idea is to find orthogonal P and Q such that

$$P^\top \begin{bmatrix} b & AQ \end{bmatrix} = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline\hline 0 & 0 & A_{22} \end{array}\right]. \tag{13}$$

The block A_{11} is of full column rank, has simple singular values and b_1 has nonzero projections onto the left singular vectors of A_{11}. These properties guarantee that the subproblem $A_{11}\,x_1 \;\approx\; b_1$ has minimal dimensions and contains all necessary and sufficient information for solving the original problem $A\,x \;\approx\; b$. All irrelevant and redundant information is contained in A_{22} .

Low-rank approximation. TLS problems aim at approximate solutions of overdetermined linear systems of equations $AX \;\approx\; B$. Typical application of TLS methods, however, are problems for data approximation by linear models. Such problems are mathematically equivalent to low rank approximation, which in turn is *not* equivalent to the $AX \approx B$ problem [Markovsky, 2007]. This suggests that from a data modeling point of view, a low rank approximation is a better framework than the solution of an overdetermined linear system of equations. This viewpoint of the TLS data modeling approach is presented in [Markovsky and Van Huffel, 2007a].

Application of STLS in system identification and model reduction is described in [Roorda and Heij, 1995, Markovsky et al., 2005, 2006]. Further applications of STLS include the shape from moments problem [Schuermans et al., 2006], approximate factorization and greatest common divisor computation in computer algebra [Botting, 2004], and image deblurring [Pruessner and O'Leary, 2003, Mastronardi et al., 2005]. The WTLS problem has applications in chemometrics [Wentzell et al., 1997, Schuermans et al., 2005] and machine learning [Srebro, 2004].

Applications of RTLS. RTLS formulations, including weighted and structured generalizations, have been used in various ill-posed problems. A notorious inverse problem—blind deconvolution of one- or two-dimensional data—has received special attention. Restoring one-dimensional signals from noisy measurements of both the point-spread function and the observed data has been addressed by Fan [1992], Younan and Fan [1998] as a regularized structured total least squares problem. A two-dimensional generalization has been used for image restoration in Mesarović et al. [1995]. Interesting structured regularized problem formulations and efficient algorithms for image deblurring are analyzed in Kamm and Nagy [1997], Ng et al. [2000], Chen et al. [2000], Pruessner and O'Leary [2003], Mastronardi et al. [2004, 2005], Kalsi and O'Leary [2006], Fu et al. [2006]. RTLS has also been used in image reconstruction of electrical capacitance tomography [Lei et al., 2008].

Applications of TTLS. TTLS has successfully been applied to biomedical inverse problems such as the reconstruction of epicardial potentials from body surface potentials [Shou et al., 2008] and imaging by ultrasound inverse scattering [Liu et al., 2003]. TTLS is also used as an alternative to ridge regression in the estimation step of the regularized expectation-maximization algorithm for the analysis of incomplete climate data [Schneider, 2001].

Acknowledgments

Dr. Ivan Markovsky acknowledges PinView (Personal Information Navigator adapting through VIEWing), an EU FP7 funded Collaborative Project 216529.
Dr. Diana M. Sima is a postdoctoral fellow of the Fund for Scientific Research-Flanders. Dr. Sabine Van Huffel is a full professor at the Katholieke Universiteit Leuven, Belgium. This research is supported by

- Research Council KUL: GOA MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC)
- Flemish Government: Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011)

References

T. J. Abatzoglou, I. M. Mendel, and G. A. Harada. The constrained total least-squares technique and its applications to harmonic superresolution. *IEEE Trans. Signal Process.*, 39(5):1070–1087, 1991.

A. Beck. The matrix-restricted total least-squares problem. *Signal Process.*, 87(10):2303–2312, 2007.

A. Beck and A. Ben-Tal. A global solution for the structured total least squares problem with block circulant matrices. *SIAM J. Matrix Anal. Appl.*, 27(1):238–255, 2006a.

A. Beck and A. Ben-Tal. On the solution of the Tikhonov regularization of the regularized total least squares problem. *SIAM J. Optim.*, 17:98–118, 2006b.

A. Beck, A. Ben-Tal, and M. Teboulle. 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(2):425–445, 2006.

B. Botting. Structured total least squares for approximate polynomial operations. Master's thesis, School of Computer Science, Univ. of Waterloo, 2004.

W. Chen, M. Chen, and J. Zhou. Adaptively regularized constrained total least-squares image restoration. *IEEE Trans. Image Process.*, 9(4):588–596, Apr 2000.

B. De Moor. Structured total least squares and L_2 approximation problems. *Linear Algebra Appl.*, 188–189:163–207, 1993.

X. Fan. *The constrained total least squares with regularization and its use in ill-conditioned signal restoration*. PhD thesis, Elec. Comput. Eng. Dept., Mississippi State Univ., 1992.

R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary. Regularization by truncated total least squares. *SIAM J. Sci. Comput.*, 18(1):1223–1241, 1997.

H. Fu, M. K. Ng, and J. L. Barlow. Structured total least squares for color image restoration. *SIAM J. Sci. Comput.*, 28(3):1100–1119, 2006.

W. Gander. Least squares with a quadratic constraint. *Numer. Math.*, 36:291–307, 1981.

G. Golub and V. Pereyra. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl.*, 19:1–26, 2003.

G. H. Golub, P. C. Hansen, and D. P. O'Leary. Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl.*, 21(1):185–194, 1999.

H. Guo and R. Renaut. Parallel variable distribution for total least squares. *Numer. Linear. Algebra Appl.*, 12(9):859–876, 2005.

P.-C. Hansen. Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank. *SIAM J. Sci. Stat. Comput.*, 11(3):503–518, 1990.

A. Kalsi and D. P. O'Leary. Algorithms for structured total least squares problems with applications to blind image deblurring. *J. Res. Natl. Inst. Stan.*, 111 (2):113–119, 2006.

J. Kamm and J. G. Nagy. Least squares and total least squares methods in image restoration. In *WNAA '96: Proc. 1st Intl. Workshop Num. Anal. and Its Appl.*, pages 212–219, London, UK, 1997. Springer-Verlag.

J. Lampe and H. Voss. On a quadratic eigenproblem occurring in regularized total least squares. *Comput. Stat. Data Anal.*, 52(2):1090–1102, 2007.

J. Lampe and H. Voss. A fast algorithm for solving regularized total least squares problems. *Electron. Trans. Numer. Anal.*, 31:12–24, 2008.

J. Lei, S. Liu, Z. Li, H. Schlaberg, and M. Sun. An image reconstruction algorithm based on the regularized total least squares method for electrical capacitance tomography. *Flow Meas. Instrum.*, 19(6):325–330, 2008.

C. Liu, Y. Wang, and P. A. Heng. A comparison of truncated total least squares with tikhonov regularization in imaging by ultrasound inverse scattering. *Phys. Med. Biol.*, 48(15):2437–2451, 2003.

S. Lu, S. V. Pereverzev, and U. Tautenhahn. Dual regularized total least squares and multi-parameter regularization. *Comput. Meth. Appl. Math.*, 8(3): 253–262, 2008.

J. Manton, R. Mahony, and Y. Hua. The geometry of weighted low-rank approximations. *IEEE Trans. Signal Process.*, 51(2):500–514, 2003.

I. Markovsky. Structured low-rank approximation and its applications. *Automatica*, 44(4):891–909, 2007.

I. Markovsky and S. Van Huffel. Overview of total least squares methods. *Signal Process.*, 87:2283–2302, 2007a.

I. Markovsky and S. Van Huffel. Left vs right representations for solving weighted low rank approximation problems. *Linear Algebra Appl.*, 422:540–552, 2007b.

I. Markovsky, J. C. Willems, S. Van Huffel, B. D. Moor, and R. Pintelon. Application of structured total least squares for system identification and model reduction. *IEEE Trans. Automat. Contr.*, 50(10):1490–1500, 2005.

I. Markovsky, J. C. Willems, S. Van Huffel, and B. De Moor. *Exact and Approximate Modeling of Linear Systems: A Behavioral Approach*. SIAM, March 2006.

N. Mastronardi, P. Lemmerling, A. Kalsi, D. O'Leary, and S. Van Huffel. Implementation of the regularized structured total least squares algorithms for blind image deblurring. *Linear Algebra Appl.*, 391:203–221, 2004.

N. Mastronardi, P. Lemmerling, and S. Van Huffel. Fast regularized structured total least squares problems for solving the basic deconvolution problem. *Numer. Linear. Algebra Appl.*, 12(2–3):201–209, 2005.

V. Mesarović, N. Galatsanos, and A. Katsaggelos. Regularized constrained total least squares image restoration. *IEEE Trans. Image Process.*, 4(8):1096–1108, 1995.

M. K. Ng, R. J. Plemmons, and F. Pimentel. A new approach to constrained total least squares image restoration. *Linear Algebra Appl.*, 316(1–3):237–258, 2000.

C. C. Paige and Z. Strakoš. Core problems in linear algebraic systems. *Comput. Stat. Data Anal.*, 27(3):861–875, 2005.

A. Pruessner and D. P. O'Leary. Blind deconvolution using a regularized structured total least norm approach. *SIAM J. Matrix Anal. Appl.*, 24:1018–1037, 2003.

R. Renaut and H. Guo. Efficient algorithms for solution of regularized total least squares. *SIAM J. Matrix Anal. Appl.*, 26(2):457–476, 2005.

B. Roorda and C. Heij. Global total least squares modeling of multivariate time series. *IEEE Trans. Automat. Contr.*, 40(1):50–63, 1995.

T. Schneider. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate*, 14:853–871, 2001.

M. Schuermans, I. Markovsky, P. Wentzell, and S. Van Huffel. On the equivalence between total least squares and maximum likelihood PCA. *Anal. Chim. Acta*, 544:254–267, 2005.

M. Schuermans, P. Lemmerling, L. D. Lathauwer, and S. Van Huffel. The use of total least squares data fitting in the shape from moments problem. *Signal Process.*, 86:1109–1115, 2006.

G. Shou, L. Xia, M. Jiang, Q. Wei, F. Liu, and S. Crozier. Truncated total least squares: A new regularization method for the solution of ECG inverse problems. *IEEE Trans. Biomed. Eng.*, 55(4):1327–1335, 2008.

D. M. Sima and S. Van Huffel. Level choice in truncated total least squares. *Comput. Stat. Data Anal.*, 52(2):1103–1118, 2007.

D. M. Sima, S. Van Huffel, and G. H. Golub. Regularized total least squares based on quadratic eigenvalue problem solvers. *BIT*, 44(4):793–812, December 2004.

N. Srebro. *Learning with Matrix Factorizations*. PhD thesis, MIT, 2004.

A. N. Tikhonov and V. Arsenin. *Solutions of Ill-Posed Problems*. Winston & Sons, Washington, DC, USA, 1977.

S. Van Huffel, editor. *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*. SIAM, Philadelphia, 1997.

S. Van Huffel. Total least squares and errors-in-variables modeling: Bridging the gap between statistics, computational mathematics and engineering. In J. Antoch, editor, *COMPSTAT Proceedings in Computational Statistics*, pages 539–555. Physika-Verlag, Heidelberg, 2004.

S. Van Huffel and P. Lemmerling, editors. *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*. Kluwer, 2002.

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

S. Van Huffel and H. Zha. The total least squares problem. In C. R. Rao, editor, *Handbook of Statistics: Computational Statistics*, volume 9, pages 377–408. Elsevier Science Publishers B.V., Amsterdam, 1993.

S. Van Huffel, C.-L. Cheng, N. Mastronardi, C. Paige, and A. Kukush. Editorial: Total least squares and errors-in-variables modeling. *Comput. Stat. Data Anal.*, 52:1076–1079, 2007a.

S. Van Huffel, I. Markovsky, R. J. Vaccaro, and T. Söderström. Guest editorial: Total least squares and errors-in-variables modeling. *Signal Process.*, 87(10): 2281–2282, October 2007b.

P. Wentzell, D. Andrews, D. Hamilton, K. Faber, and B. Kowalski. Maximum likelihood principal component analysis. *J. Chemometr.*, 11:339–366, 1997.

N. H. Younan and X. Fan. Signal restoration via the regularized constrained total least squares. *Signal Process.*, 71:85–93, 1998.

W. Zhu, Y. Wang, N. P. Galatsanos, and J. Zhang. Regularized total least squares approach for nonconvolutional linear inverse problems. *IEEE Trans. Image Process.*, 8(11):1657–1661, 1999.