Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 544 (2005) 254–267
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

# On the equivalence between total least squares and maximum likelihood PCA

M. Schuermans[a,*], I. Markovsky[a], Peter D. Wentzell[b], S. Van Huffel[a]

[a] *KULeuven, ESAT-SCD, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium*
[b] *Trace Analysis Research Centre, Department of Chemistry, Dalhousie University, Halifax, NS, Canada B3H 4J3*

Available online 26 January 2005

**Abstract**

The maximum likelihood PCA (MLPCA) method has been devised in chemometrics as a generalization of the well-known PCA method in order to derive consistent estimators in the presence of errors with known error distribution. For similar reasons, the total least squares (TLS) method has been generalized in the field of computational mathematics and engineering to maintain consistency of the parameter estimates in linear models with measurement errors of known distribution. The basic motivation for TLS is the following. Let a set of multidimensional data points (vectors) be given. How can one obtain a linear model that explains these data? The idea is to modify all data points in such a way that some norm of the modification is minimized subject to the constraint that the modified vectors satisfy a linear relation. Although the name "total least squares" appeared in the literature only 25 years ago, this method of fitting is certainly not new and has a long history in the statistical literature, where the method is known as "orthogonal regression", "errors-in-variables regression" or "measurement error modeling". The purpose of this paper is to explore the tight equivalences between MLPCA and element-wise weighted TLS (EW-TLS). Despite their seemingly different problem formulation, it is shown that both methods can be reduced to the same mathematical kernel problem, i.e. finding the closest (in a certain sense) weighted low rank matrix approximation where the weight is derived from the distribution of the errors in the data. Different solution approaches, as used in MLPCA and EW-TLS, are discussed. In particular, we will discuss the weighted low rank approximation (WLRA), the MLPCA, the EW-TLS and the generalized TLS (GTLS) problems. These four approaches tackle an equivalent weighted low rank approximation problem, but different algorithms are used to come up with the best approximation matrix. We will compare their computation times on chemical data and discuss their convergence behavior.
© 2004 Elsevier B.V. All rights reserved.

*Keywords:* TLS; MLPCA; Rank reduction; Measurement errors

## 1. Introduction

### *1.1. Total least squares: problem formulation, algorithms and extensions*

The total least squares (TLS) method is one of several linear parameter estimation techniques that has been devised to compensate for data errors. The univariate line fitting problem was already discussed since 1877 [1]. More recently, the TLS approach to fitting has also stimulated interests outside statistics. One of the main reasons for its popularity is the availability of efficient and numerically robust algorithms in which the singular value decomposition (SVD) plays a prominent role [2]. Another reason is the fact that TLS is an application oriented procedure. It is suited for situations in which all data are corrupted by noise, which is almost always the case in engineering applications. In this sense, TLS and errors-in-variables (EIV) modeling are a powerful extension of classical least squares and ordinary regression, which corresponds only to a partial modification of the data. A comprehensive description of the state of the art on TLS from its conception up to the summer of 1990 and its use in parameter estimation has been presented in [3]. While the latter book is entirely devoted to TLS, a second [4] and third book [5] present the progress in TLS and in the broader field of errors-in-variables modeling respectively from 1990 till 1996 and from 1996 till 2001.

* Corresponding author. Tel.: +32-16321143.
*E-mail address:* mieke.schuermans@esat.kuleuven.ac.be (M. Schuermans).

0304-4068/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.12.059

The problem of *linear parameter estimation* arises in a broad class of scientific disciplines such as signal processing, automatic control, system theory and in general engineering, statistics, physics, economics, biology, medicine, etc. It starts from a model described by a linear equation:

$$\xi_1 b_1 + \cdots + \xi_n b_n = \eta \tag{1}$$

where $\xi_1, \ldots, \xi_n$ and $\eta$ denote the variables and $b = [b_1, \ldots, b_n]^\top \in \mathbb{R}^n$ plays the role of a parameter vector that characterizes the specific system. A basic problem of applied mathematics is to determine an estimate of the *true* but *unknown* parameters from given measurements of the variables. In general, $\eta$ can be a vector of dimension $d > 1$. In this case, the parameters form a matrix $B$ and more than one (independent) linear relationship between $\eta$ and the variables $\xi_i$ has to be estimated. This gives rise to an *overdetermined* set of $m$ linear equations ($m > n$):

$$XB \approx Y \tag{2}$$

where the $i$th row of data matrix $X \in \mathbb{R}^{m\times n}$ and vector $Y \in \mathbb{R}^{m\times d}$ contain respectively the measurements of the variables $\xi_1, \ldots, \xi_n$ and $\eta$.

In the classical least squares approach, as commonly used in ordinary regression, the measurements $X$ of the variables $\xi_i$ are assumed to be free of error and hence, all errors are confined to the observations $Y$. However, this assumption is frequently unrealistic: sampling errors, human errors, modeling errors and instrument errors may imply inaccuracies of the data matrix $X$ as well. One way to take errors in $X$ into account is to introduce *perturbations also in X*. Therefore, the following TLS problem was introduced in the field of computational mathematics [2,6] ($R(X)$ denotes the range of $X$ and $\|X\|_F$ its Frobenius norm [7]):

**Definition 1** (Total least squares problem). Given an overdetermined set of $m$ linear equations $XB \approx Y$ in $n \times d$ unknowns B. The total least squares problem seeks to

$$\min_{\widehat{\Delta X}, \widehat{\Delta Y}, \hat{B}} \|[\,\widehat{\Delta X}\ \widehat{\Delta Y}\,]\|_F \text{ subject to } (X - \widehat{\Delta X})\hat{B} = Y - \widehat{\Delta Y} \tag{3}$$

$\hat{B}$ is called a TLS solution and $[\,\widehat{\Delta X}\ \widehat{\Delta Y}\,]$ the corresponding TLS correction.

In most multivariate problems, the TLS problem (3) has a unique solution which can be obtained from a simple scaling of the $d$ right singular vectors of $[\,X\ Y\,]$ corresponding to its $d$ smallest singular values. Extensions of this basic TLS problem to problems in which the TLS solution is no longer unique (called minimum norm TLS since then the solution with minimal norm is selected) or fails to have a solution altogether (called non-generic TLS) and to mixed LS-TLS problems that assume some of the columns of $X$ to be error-free, are considered in detail in [3]. In addition, it is shown how to speed up the TLS computations directly by computing the SVD only partially or iteratively if a good starting vector is available. More recent advances, e.g. recursive TLS algorithms, neural based TLS algorithms, rank-revealing TLS algorithms, regularized TLS algorithms, TLS algorithms for large scale problems, etc., are reviewed in [4,5].

Under specific conditions, the TLS solution, as introduced in numerical analysis, computes optimal parameter estimates in models with *only measurement error*, referred to as classical *EIV* models. These models are characterized by the fact that the true values of the observed variables satisfy one or more unknown but exact linear relations of the form (1). In particular, we define [8–10] the following.

**Definition 2** (Multivariate EIV regression model). Assume that the $m$ measurements in $X, Y$ are related to $n \times d$ unknowns $B$ by:

$$\Xi B = \eta, \quad X = \Xi + \Delta X \text{ and } Y = \eta + \Delta Y \tag{4}$$

where $\Delta X$, $\Delta Y$ represent the measurement errors and all rows of $[\,\Delta X\ \Delta Y\,]$ are i.i.d. with zero mean and covariance matrix $\mathcal{C}$, known up to a scalar multiple $\sigma_v^2$.

If additionally $\mathcal{C} = \sigma_v^2 I$ is assumed with $I$ the identity matrix (i.e. $\Delta x_{ij}$ and $\Delta y_{ij}$ are uncorrelated random variables with equal variance) and $\lim_{m\to\infty} \frac{1}{m} \Xi^\top \Xi$ exists and is positive definite, then it can be proven [9,6] that the TLS solution $\hat{B}$ of $XB \approx Y$ estimates the *true* parameter values $B$, given by $(\Xi^\top \Xi)^{-1} \Xi^\top \eta$, *consistently*, i.e. $\hat{B}$ converges to $B$ as $m \to \infty$. This TLS property does not depend on any assumed distributional form of the errors. If additionally the errors are normal, then TLS computes the ML estimate. It should be noted that the TLS correction $[\,\widehat{\Delta X}\ \widehat{\Delta Y}\,]$, being of rank $\leq d$, cannot be considered as an appropriate estimator for the true measurement errors $\Delta X$ and $\Delta Y$, added to the data [2,3]. Note also that the LS estimates are inconsistent in this case. In these cases, TLS gives better estimates than does LS, as confirmed by simulations [3]. This situation may occur far more often in practice than is recognized. It is very common in agricultural, medical and economic science, in humanities, business and many other data analysis situations. Hence TLS should be a quite useful tool to data analysts. In fact, the key role and importance of LS in regression analysis is the same as that of TLS in EIV regression. Nevertheless, a lot of confusion exists in the fields of numerical analysis and statistics about the principle of TLS and its relation to EIV modeling. Roughly speaking, TLS is a special case of EIV estimation and, as such, TLS is reduced to a method in statistics but, on the other hand, TLS appears in many other fields, where mainly the data modification idea is used and explained from a geometric point of view, independently from its statistical interpretation.

The statistical model that corresponds to the basic TLS approach is the no-equation-error EIV regression model with the restrictive condition that the measurement errors on the data are i.i.d. with zero mean and common error covari-

ance matrix, equal to the identity matrix up to an unknown scalar. Most published TLS algorithms just handle this case while other more useful EIV regression estimators did not receive enough attention in computational mathematics. To relax these restrictions, several extensions of the TLS problem have been investigated. In particular, the *mixed LS-TLS* problem formulation allows extension of the consistency of the TLS estimator in EIV models, where some of the variables $\xi_i$ are measured without error. The *data least squares* problem refers to the special case in which all variables except $\eta$ are measured with error and was introduced in the field of signal processing by DeGroat and Dowling [11] in the mid-1990s. Whenever the errors are independent but *unequally* sized, *weighted TLS* problems should be considered using appropriate diagonal scaling matrices in order to maintain consistency. If, additionally, the errors are also correlated, then the *generalized TLS* (GTLS) [12] problem formulation (defined in the next section) allows extension of the consistency of the TLS estimator in EIV models, provided the corresponding error covariance matrix $\mathcal{C}$ is known up to a factor of proportionality (see Definition 2).

More general problem formulations, such as *restricted TLS*, which also allow the incorporation of equality constraints, have been proposed, as well as equivalent problem formulations using other $L_p$ norms and resulting in the so-called *Total* $L_p$ *approximations* (see [3] for references). The latter problems proved to be useful in the presence of outliers. Robustness of the TLS solution is also improved by adding regularization, resulting in the *regularized TLS* methods [5,13,14]. In addition, various types of bounded uncertainties have been proposed in order to improve robustness of the estimators under various noise conditions and algorithms are outlined [4,5].

Furthermore, *constrained TLS* problems have been formulated. Arun [15] addressed the unitarily constrained TLS problem, i.e., $XB \approx Y$, subject to the constraint that the solution matrix $B$ should be unitary. He proved that this solution is the same as the solution to the orthogonal Procrustes problem [7, p. 582]. Abatzoglou et al. [16] considered yet another constrained TLS problem, which extends the classical TLS problem (3) to the case where the errors $[\,\Delta X\ \Delta Y\,]$ in the data $[\,X\ Y\,]$ are algebraically related. However, if there is a linear dependence among the error entries in $[\,\Delta X\ \Delta Y\,]$, then the TLS solution no longer has optimal statistical properties (e.g. maximum likelihood in case of normality). This happens, for instance, in dynamic system modeling when we try to estimate the impulse response of a system from its input and output by discrete deconvolution. In these so-called *structured TLS* problems, the data matrix $[\,X\ Y\,]$ is structured, typically block Toeplitz or Hankel. In order to preserve maximum likelihood properties and consistency of the solution [16,17], the TLS problem formulation, given in Definition 1, must be extended with the additional constraint that any (affine) structure of $X$ or $[\,X\ Y\,]$ must be preserved in $\widehat{\Delta X}$ or $[\,\widehat{\Delta X}\ \widehat{\Delta Y}\,]$, where $\widehat{\Delta X}$ and $\widehat{\Delta Y}$ are chosen to minimize the error in the discrete $L_1$, $L_2$ and $L_\infty$ norm. For $L_2$ norm minimization, various computational algorithms have been presented, as surveyed in [4,5], and shown to reduce the computation time by exploiting the matrix structure in the computations. In addition, it is shown how to extend the problem and solve it, if latency or equation errors are included. Recently, robustness of the structured TLS solution has been improved by adding regularization, see e.g. [18].

Yet, another important extension is the *element-wise weighted TLS* (EW-TLS) estimator [19,20] (see definition in the next section). EW-TLS computes consistent estimates in linear EIV models, where the measurement errors are element-wise differently sized or, more generally, where the corresponding error covariance matrices may differ from row to row, provided the total covariance structure of the errors is known up to a scalar factor. Some of the variables are allowed to be exactly known (observable) [5,21]. Mild conditions for weak consistency of the EW-TLS estimator have been given and an iterative procedure to compute it has been proposed.

### 1.2. Total least squares in chemometrics: state-of-the-art

Although TLS has already been applied in chemometrics since 1984 [22], its concept and properties are not well known in chemometrics and in fact a lot of confusion exists about its proper use and relation to well-known chemometric techniques such as principal component analysis (PCA), partial least squares (PLS) and other EIV approaches such as latent root regression (LRR).

Several authors, such as Burnham et al. [23], de Jong [24,25] and Phatak [4,26,27], have described the close relationships between a variety of multivariate regression techniques, including TLS, mainly from a theoretical point of view. Although TLS should be primarily used for parameter estimation, it can be used for prediction provided the solution is appropriately computed [22,28]. In [22], the prediction accuracy and stability of the TLS estimator is compared to that of ordinary LS, PC regression (PCR), ridge regression and shrunken estimators in a chemical example in multivariate regression in the presence of multicollinearities among the independent variables. In many chemical applications of multivariate regression like the present example of relationships between chemical structure and biological activity, the predictive properties of the model are of prime importance and the regression estimates therefore often need to be stabilized. The example [22] shows that minimum norm TLS gives a solution to the multivariate regression problem which is stabilized in comparison with the LS solution and which has (at least in the example investigated) minimal prediction error. Stabilization is performed by reducing the matrix $[\,X\ y\,]$ to a matrix of much smaller rank.

More generally, the usefulness of an EIV model for providing insights into the multivariate calibration problem [29] was demonstrated by Thomas [4,30]. Predicting new concentrations of a multicomponent sample from the associated new spectra can be viewed as parameter estimation in an EIV

framework with maximum likelihood properties. Moreover, using the responses (spectra) from the available members of the prediction set together improves the accuracy of the new concentrations compared to estimation based on only one response of the current member of the prediction set.

Another EIV approach to multivariate calibration and multivariate regression, two kernel problems in analytical chemistry, was developed by Wentzell et al. [28], who generalized the PCA method to maximum likelihood PCA (MLPCA) [31,32]. MLPCA is an analog to PCA that incorporates information about measurement errors to develop PCA models that are optimal in a maximum likelihood sense. A globally convergent algorithm based on an alternating regression method has been derived. Generalization of the algorithm allows its adaptation to cases of correlated errors provided the error covariance matrix is known. Simplifications to the original algorithm apply when errors are correlated only within the rows of a data matrix and when all of these row covariance matrices are equal. Simulated and experimental data for three-component mixtures show that inclusion of error covariance information via MLPCA outperforms PCA when the true error covariance matrix is available. Using MLPCA, two approaches to multivariate calibration have been described that allow information on measurement uncertainties to be included in the calibration process in a statistically meaningful way. These methods, referred to as maximum likelihood PCR (MLPCR) and maximum likelihood LRR (MLLRR) are based on principles of ML parameter estimation and generalize PCR, which has been widely used in chemistry, and LRR [33,34], which has been almost ignored in this field. Whereas MLPCR is based on the decomposition of the calibration matrix $X$ by MLPCA, MLLRR decomposes the augmented data matrix $[\,X\ Y\,]$ using MLPCA. In multivariate calibration, this is the original calibration matrix of response variables $X$ augmented by the corresponding concentration vectors $Y$. (ML)LRR is equivalent to (EW)TLS in the sense that it computes the best lower rank approximation $\widehat{D} = [\,\hat{X}\ \hat{Y}\,]$ of the augmented data matrix $D = [\,X\ Y\,]$ in a (weighted) least squares sense. If $\hat{X} \in R(\hat{Y})$, TLS computes the minimum norm solution which equals the minimum norm LS solution and hence TLS is equivalent to LRR, as reported in [28]. However, if only non-predictive multicollinearities (i.e. linear dependencies among the columns of $X$ only) have been removed via the rank reduction, then $\hat{X} \notin R(\hat{Y})$ and the TLS problem is called non-generic. In this case, the LRR solution differs from the non-generic TLS solution, as shown in [3]. After removing the non-predictive multicollinearities, the LRR estimate is found by minimizing a sum of squared residuals, while the non-generic TLS solution is found by minimizing the corrections applied to the data such that the corrected data satisfy the constraint $\hat{X} \in R(\hat{Y})$. By using estimates of the measurement error variance, MLPCR and MLLRR are able to extract the optimum amount of information from each measurement and, thereby, outperform conventional multivariate calibration methods such as PCR and PLS when there is non-uniform error structure. Using simulated and experimental data of three-component mixtures, it is shown that the ML methods outperform PCR and PLS in case of non-uniform errors. MLLRR generally performed better than MLPCR, but in most cases the improvement was marginal.

Finally, it should be noted that structured TLS (STLS) clearly differs from dynamic PCA [35]. Although both methods focus on modeling dynamic systems in an EIV context, dynamic PCA is based on computing the truncated SVD of an input/output data matrix and as such does not preserve the matrix structure during rank reduction in contrast with STLS and is hence not optimal.

### *1.3. The use of TLS and extensions in multivariate (ML)PCR*

From the previous subsections it should be clear that TLS and its extensions can be used in any multivariate (ML) regression/calibration problem $XB \approx Y$ by computing the best lower rank approximation $\hat{D} = [\,\hat{X}\ \hat{Y}\,]$ of the augmented data matrix $D = [\,X\ Y\,]$ according to a weighted norm. As such, (EW)TLS is equivalent to (ML)LRR in problem formulation. However, (EW)TLS can also be used in any multivariate (ML)PCR problem $XB \approx Y$ where an MLPCA of $X$ is sought prior to the computation of $B$. Indeed, assume that the true pseudorank of $X$ is $p$ and the measurement errors on the elements of $X$ are i.i.d. normal, then the optimal rank $p$ matrix approximation $\hat{X} \in \mathbb{R}^{m\times N}$ in an ML sense is obtained by solving the TLS problem $X_1 B \approx X_2$ where $X$ is split up in two parts $X = [\,X_1\ X_2\,]$ such that $X_1$ has $p$ columns and rank $p$. The TLS approach then computes minimal (with respect to the Frobenius norm) corrections $\widehat{\Delta X} = [\,X_1 - \hat{X}_1\ X_2 - \hat{X}_2\,]$ such that $\hat{X} = [\,\hat{X}_1\ \hat{X}_2\,]$ is of rank $p$. In other words, TLS projects the data $X$ into a lower dimensional (of dimension $p$) subspace $R(\hat{X})$ such that exactly $N - p$ independent linear relations are imposed between $\hat{X}_1$ and $\hat{X}_2$. Equivalently, for correlated normally distributed errors where the error covariance matrix $W_i$ of the errors in each row of the data matrix $X$ is known, an optimal lower rank approximation in an ML sense is obtained by minimizing a weighted sum of squared corrections. If all $W_i$ are equal to each other, EW-TLS reduces to GTLS. As such, TLS and EW-TLS algorithms can be applied to compute the PCA and MLPCA of a matrix $D = X$, as used in PCR and MLPCR, or $D = [\,X\ Y\,]$, as used in LRR and MLLRR. Other methods exist in the literature to tackle the weighted low rank matrix approximation problem, such as the weighted low rank approximation (WLRA) method [36]. It is the goal of this paper to compare the performance of the algorithms used for solving TLS problems and its extensions with the algorithms used in chemometrics for computing PCA and MLPCA. Also, the WLRA algorithm will be considered in this comparison.

The paper is organized as follows. In Sections 2 and 3, we will discuss the WLRA, the MLPCA, the EW-TLS and the GTLS problems. These four approaches tackle an equivalent weighted low rank approximation problem, but different algorithms are used to come up with the best approximation

matrix $\hat{D}$ of a given data matrix $D$. In Section 4 we will compare their computation times on simulated as well as real-life chemical data and discuss their convergence behavior. Conclusions are made in Section 5.

## 2. Weighted low rank matrix approximation: problem formulation

When the measurement noise is centered, normal and independent identically distributed, the optimal closeness norm is the Frobenius norm, $\|\cdot\|_F$. This is used in the well-known TLS and the PCA methods. Nevertheless, when the measurement errors are not identically distributed the Frobenius norm is no longer optimal and a weighted norm is needed instead. The weighting is related to the noise covariance matrix. In this case a closed-form solution is no longer available and is only implicitly defined via a non-convex optimization problem, which is the case in this paper.

Mathematically, we will consider the following weighted low rank matrix approximation problem:

$$\min_{\hat{D}} \operatorname{vec}^\top(D-\hat{D})W^{-1}\operatorname{vec}(D-\hat{D}) \quad \text{s.t. } \operatorname{rank}(\hat{D}) \le p \quad (5)$$

with $D \in \mathbb{R}^{m\times N}$, $m \le N$, the noisy data matrix, $p < m$, $\widehat{\Delta D} = D - \hat{D}$ the estimated measurement noise and $W$ the covariance matrix of $\operatorname{vec}(\widehat{\Delta D})$ where $\operatorname{vec}(\widehat{\Delta D})$ stands for the vectorized form of $\widehat{\Delta D}$, i.e., a vector constructed by stacking the consecutive columns of $\widehat{\Delta D}$ in one vector.

We will discuss the WLRA, the MLPCA, the EW-TLS and the GTLS method to tackle this matrix weighted low rank approximation problem. These methods differ in the representation of the rank constraint $\operatorname{rank}(\hat{D}) \le p$ in problem (5) and in the applied optimization technique in order to solve problem (5). In this section a short description is given of the different problem formulations and the different approaches for these methods. In the next section their different algorithms will be discussed.

The different representations of the rank constraint are the following:

$\mathcal{C}1$: $\hat{D} = TP^\top$, where $T \in \mathbb{R}^{m\times p}$ and $P \in \mathbb{R}^{N\times p}$;
$\mathcal{C}2$: $\hat{D}\mathcal{N} = 0$, where $\mathcal{N} \in \mathbb{R}^{N\times(N-p)}$ and $\mathcal{N}^\top\mathcal{N} = I$;
$\mathcal{C}3$: $\hat{D}\begin{bmatrix} B \\ -I_{N-p}\end{bmatrix} = 0$, where $B \in \mathbb{R}^{p\times(N-p)}$.

$\mathcal{C}1$ and $\mathcal{C}2$ are equivalent. Representations $\mathcal{C}2$ and $\mathcal{C}3$ only differ in the way $\mathcal{N}$ is restricted to be full rank. Representation $\mathcal{C}2$ imposes the additional constraint $\mathcal{N}^\top\mathcal{N} = I$, while representation $\mathcal{C}3$ sets the lower $(N-p)\times(N-p)$ block of $\mathcal{N}$ to $-I$. The constraint $\mathcal{N}^\top\mathcal{N} = I$ has the advantage that it avoids the so-called non-generic cases, but it is more difficult to deal with. Indeed, by using $\mathcal{C}3$ for representing the rank constraint, the condensed problem is an unconstrained optimization problem, while using $\mathcal{C}2$ the condensed problem remains to a constrained optimization problem since the constraint $\mathcal{N}^\top\mathcal{N} = I$ cannot be substituted in the object function.

Any of the representations $\mathcal{C}1$ – $\mathcal{C}3$ for the rank constraint is bilinear. To handle such a bilinear problem, different optimization techniques can be used:

$\mathcal{T}1$: applying an alternating least squares algorithm;
$\mathcal{T}2$: eliminating one of the variables and obtaining an equivalent problem in the other variable;
$\mathcal{T}3$: linearizing the constraint and solving iteratively equality-constrained least squares problems.

Now, we can formulate the WLRA [36], the MLPCA [31], the EW-TLS [19] and the GTLS [12] problem:

**WLRA** : $\min_{\widehat{\Delta D},\mathcal{N}} \operatorname{vec}^\top(\widehat{\Delta D})W^{-1}\operatorname{vec}(\widehat{\Delta D})$ s.t. $(D-\widehat{\Delta D})\mathcal{N} = 0$, where $\mathcal{N} \in \mathbb{R}^{N\times(m-p)}$ and $\mathcal{N}^\top\mathcal{N} = I$.

**MLPCA** : $\min_{T,P} \operatorname{vec}^\top(\widehat{\Delta D})W^{-1}\operatorname{vec}(\widehat{\Delta D})$ s.t. $\widehat{\Delta D} = D - TP^\top$, with $T \in \mathbb{R}^{m\times p}$ and $P \in \mathbb{R}^{N\times p}$.

**EW-TLS** : $\min_{\hat{B},\widehat{\Delta D}} \sum_{i=1}^{m} \operatorname{vec}^\top(\widehat{\Delta D}_i)W_i^{-1}\operatorname{vec}(\widehat{\Delta D}_i)$ s.t. $(D-\widehat{\Delta D})\begin{bmatrix}\hat{B}\\ -I\end{bmatrix} = 0$ with $\widehat{\Delta D}_i \in \mathbb{R}^N$ the $i$th row of $\widehat{\Delta D}$ and $W_i$ the $i$th weighting matrix defined as the covariance matrix of the errors in the $i$th row of the data matrix $D$.

**GTLS** : $\min_{\hat{B},\widehat{\Delta D}_2} \sum_{i=1}^{m} \operatorname{vec}^\top(\widehat{\Delta D}_{2,i})W_f^{-1}\operatorname{vec}(\widehat{\Delta D}_{2,i})$ s.t. $(D-\widehat{\Delta D})\begin{bmatrix}\hat{B}\\ -I\end{bmatrix} = 0$ with $\widehat{\Delta D}_{2,i} \in \mathbb{R}^{N_2}$ the $i$th row of $\widehat{\Delta D}_2$, $\widehat{\Delta D} = [0\,\widehat{\Delta D}_2] \in \mathbb{R}^{m\times N}$ and the weighting matrix $W_f$ is the covariance matrix of the errors in the $i$th row of the noisy data matrix $D_2$ assumed to be equal for all rows. $D_2 \in \mathbb{R}^{m\times N_2}$ is the noisy part of the data matrix $D \in \mathbb{R}^{m\times N}$. Some columns of $D$ can be known exactly, so $D = [\,D_1\ D_2\,] = [\,D_1\ \hat{D}_2 + \widehat{\Delta D}_2\,]$ with $D_1 \in \mathbb{R}^{m\times N_1}$ noise-free and $N_1 + N_2 = N$.

The WLRA and the MLPCA problem are the most general problem formulations. The weighting matrix $W$ can be any possible positive definite covariance matrix of the measurement noise. So, for all data sets, no matter what kind of correlation exists among their measurement errors, the WLRA and the MLPCA method can be applied. For the EW-TLS and the GTLS problem the covariance matrix $W$ has a specific block diagonal structure, $W = \operatorname{blkdiag}(W_1, \ldots, W_m)$ and $W = \operatorname{blkdiag}(W_f, \ldots, W_f)$ respectively. In terms of correlation, it means that there should be no correlation at all among the measurement errors along the columns of the data matrix. Only the rows of the measurement noise matrix can be correlated. In calibration problems [32], the rows of the data matrix are formed by individual spectra. Hence, it is reasonable to discuss the case where correlations between the measurement errors exist only along the rows. The GTLS problem is restrictive in the way that the covariance matrices of the rows of the measurement noise matrix should be the same for all rows. In this paper we consider $N_1 = 0$. Thus, all the columns of the data matrix $D$ contain measurement noise. For real-life data this is mostly the case.

Table 1
Different approaches of the WLRA, the MLPCA and the EW-TLS problem

| Approach | WLRA | MLPCA | EW-TLS |
|---|---|---|---|
| Rank constraint | $\mathcal{C}2$ | $\mathcal{C}1$ | $\mathcal{C}3$ |
| Optimization technique | $\mathcal{T}2$ | $\mathcal{T}1$ | $\mathcal{T}2$ |

Another difference between the WLRA-/MLPCA problem and the two TLS problem formulations is the approximation variable. The first two problems are matrix approximation problems to find the approximation matrix $\hat{D}$, while the latter ones are looking for a solution $\hat{B}$ of a linear model. Nevertheless, one can easily find the corresponding $\hat{B}$ from $\hat{D}$ such that $\hat{D}\begin{bmatrix}\hat{B}\\ -I\end{bmatrix} = 0$ and at any optimal solution $\hat{B}$, $\hat{D}$ from the TLS problem formulations $\hat{D}$ can be obtained from $\hat{B}$ and the weighting matrices via a closed form expression [20, Eq. (14)]. Thus except for the above-mentioned differences, the four formulated problems are equivalent. Their numerical algorithms, however, are different. This will be discussed in the next section.

In Table 1 the rank constraint representations and the optimization approaches of the four mentioned problems are summarized. Because the GTLS problem is a special case of the EW-TLS problem with $W_i = W_f$ for $i = 1, \ldots, m$, the GTLS problem is not included in the table. Fig. 3 explains the hierarchy of the four problem formulations.

## 3. Different algorithms for one problem

### 3.1. GTLS algorithm

The GTLS algorithm is derived for the special case when

$$W = \text{blkdiag}(W_f, \ldots, W_f) \text{ with } W_f \in \mathbb{R}^{N\times N}.$$

The following input/output representation of the model is used:

$$\text{rank}(\hat{D}) \le p \Longleftrightarrow \hat{X}\hat{B} = \hat{Y}, \text{ where } \hat{D} = [\,\hat{X}\ \hat{Y}\,]$$

with col dim$(\hat{X}) = p$, $\hat{X} \in \mathbb{R}^{m\times p}$ and $\hat{Y} \in \mathbb{R}^{m\times(N-p)}$. Hence, the low rank approximation problem (5) can be written as

$$\min_{\hat{B},\widehat{\Delta D}} \sum_{i=1}^{m} \text{vec}^\top(\widehat{\Delta D}_i) W_f^{-1} \text{vec}(\widehat{\Delta D}_i) \text{ s.t. } (D - \widehat{\Delta D})\begin{bmatrix}\hat{B}\\ -I\end{bmatrix} = 0 \tag{6}$$

To compute a solution of problem (6) the data matrix $D = [\,X\ Y\,]$ is scaled such that the error covariance matrix $W_f^*$ of the transformed data $D^* = [\,X^*\ Y^*\,]$ is diagonal with equal variances, i.e. $W_f^* \sim I$. The classical TLS algorithm [3] is used to solve this transformed set of equations $X^*B^* \approx Y^*$ and finally the solution of the transformed set is converted into a solution of the original set of equations $XB \approx Y$. An outline of the GTLS algorithm is given in Algorithm 1. Alternatively, the GTLS problem can also be solved by making use of the Generalized SVD [7]. This approach is recommended when $W_f$ is close to rank deficiency. For more details, the interested reader is referred to [12].

**Algorithm 1.** GTLS algorithm via rescaling the data.

1. Input: the data matrix $D \in \mathbb{R}^{m\times N}$, the error covariance matrix $W_f \in \mathbb{R}^{N\times N}$ which is equal for all rows of $D$ and a rank specification $p$.
2. Define $D = [\,X\ Y\,]$, with $X \in \mathbb{R}^{m\times p}$ and $Y \in \mathbb{R}^{m\times(N-p)}$.
3. Compute the Cholesky factorization of $W_f$ : $W_f = C^\top C$.
4. Rescale the data matrix $D$:

$$D^* = [\,X^*\ Y^*\,] = [\,X\ Y\,]C^{-1} = [\,X\ Y\,]\begin{array}{c} \\ p \\ N-p \end{array}\!\!\begin{array}{c} \begin{matrix} p & N-p \end{matrix} \\ \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix} \end{array}$$

5. Apply the classical TLS algorithm to the data $D^*$: compute a rank $p$ truncated SVD approximation $\widehat{D^*}_{\text{tls}} = U_p S_p V_p^\top$ of $D^*$, with the SVD of the matrix $D^*$ equal to $USV^\top$, $U_p$ the truncation of the matrix $U$ to $m \times p$, $S_p$ the truncation of $S$ to $p \times p$ and $V_p$ the truncation of $V$ to $N \times p$, $U_p^\top U_p = I$, $V_p^\top V_p = I$; compute the solution $\widehat{B^*}_{\text{tls}}$ of $X^*B^* \approx Y^*$:
   if $p \ge N - p$ then $\widehat{B^*}_{\text{tls}} = -V_{12}V_{22}^{-1}$, where
$$\begin{array}{c} \\ p \\ N-p \end{array}\!\!\begin{array}{c} \begin{matrix} p & N-p \end{matrix} \\ \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \end{array}$$
   else $\widehat{B^*}_{\text{tls}} = (V_{11}^\top)^{-1}V_{21}^\top$.
6. Output: $\hat{B}_{\text{gtls}} = (C_{11}\widehat{B^*}_{\text{tls}} - C_{12})C_{22}^{-1}$; $\hat{D}_{\text{gtls}} = \widehat{D^*}_{\text{tls}}C$.

**Note 1** (Special cases: ordinary TLS and weighted TLS). It is worth noting that when $W \sim I$ (i.e., the errors of the data matrix $D$ are uncorrelated and equally sized), the GTLS solution reduces to the ordinary TLS estimate. Whenever the errors are uncorrelated but unequally sized ($W$ is diagonal), the GTLS problem formulation reduces to the weighted TLS problem.

### 3.2. Alternating least squares algorithm

The alternating least squares algorithm is based on an image representation of the model

$$\text{rank}(\hat{D}) \le p \Longleftrightarrow \hat{D} = TP^\top, \text{ where } T \in \mathbb{R}^{m\times p},\ P \in \mathbb{R}^{N\times p}$$

which allows us to write the low rank approximation problem (5) equivalently as

$$\min_{T\in\mathbb{R}^{m\times p}} \min_{P\in\mathbb{R}^{N\times p}} \mathrm{vec}^\top(D-TP^\top)W^{-1}\mathrm{vec}(D-TP^\top) \quad (7)$$

The following two problems:

$$\min_{T\in\mathbb{R}^{m\times p}} \mathrm{vec}^\top(D-TP^\top)W^{-1}\mathrm{vec}(D-TP^\top) \quad (8)$$

$$\min_{P\in\mathbb{R}^{N\times p}} \mathrm{vec}^\top(D-TP^\top)W^{-1}\mathrm{vec}(D-TP^\top) \quad (9)$$

derived from (7) by fixing, respectively, $P$ and $T$ to given values are least squares problems and therefore can be solved in closed form. They can be viewed as relaxations of the non-convex problem (7) to convex problems.

The alternating least squares method is an iterative method that alternately solves (8) and (9) with respectively $P$ and $T$ fixed to the solution of the previously solved relaxation problem. The resulting alternating least squares algorithm, see Algorithm 2, is given for the case of an arbitrary positive definite covariance matrix $W$. When $W$ happens to be block diagonal or diagonal (corresponding to respectively column-wise and element-wise uncorrelated measurement errors), Algorithm 2 can be implemented more efficiently by taking into account the structure of $W$. For example, let

$$W = \mathrm{blkdiag}(W_1,\ldots,W_N), \quad \text{where } W_i\in\mathbb{R}^{m\times m}$$

and define $D=[\,d_1\ \cdots\ d_N\,]$. The solution of problem (8) can be computed efficiently in this case as follows:

$$p_i = (T^\top W_i^{-1}T)^{-1}T^\top W_i^{-1}d_i, \quad \text{for } i=1,\ldots,N,$$

where $p_i$ is the $i$th column of $P^\top$. In a similar (but not so straightforward) way, the block diagonal structure of $W$ can be taken into account for improving the computational efficiency when solving problem (9). We skip the algorithmic details.

**Algorithm 2.** Alternating least squares algorithm.

1. Input: data matrix $D\in\mathbb{R}^{m\times N}$, covariance matrix $W\in\mathbb{R}^{Nm\times Nm}$, rank specification $p$, and relative convergence tolerance $\varepsilon$.
2. Initial approximation: compute a rank $p$ GTLS approximation $\hat{D}_{\mathrm{gtls}}^\top$ of $D^\top$ with $W_f=\frac{1}{N}\sum_{i=1}^N W_i$ where $W_i$ is the submatrix of $W$ at the intersection of the rows $(i-1)m+1$ to $i\cdot m$ and the columns $(i-1)m+1$ to $i\cdot m$ (see Algorithm 1 and let $T^{(0)}=\hat{T}_{\mathrm{gtls}}$, $D^{(0)}=\hat{D}_{\mathrm{gtls}}$, $P^{(0)}=\hat{P}_{\mathrm{gtls}}$, where $\hat{P}_{\mathrm{gtls}}$ is the matrix, such that $\hat{D}_{\mathrm{gtls}}=\hat{T}_{\mathrm{gtls}}\hat{P}_{\mathrm{gtls}}^\top$.
3. $k=0$.
4. **repeat**
5. Compute the solution of (8)
$$\mathrm{vec}(P^{(k+1)\top}) = (\mathbf{M}^{(k)\top}W^{-1}\mathbf{M}^{(k)})^{-1}\mathbf{M}^{(k)\top}W^{-1}\mathrm{vec}(D),$$
where $\mathbf{M}^{(k)}=I_N\otimes T^{(k)}$.
6. $k=k+1$.
7. Compute the solution of (9)
$$\mathrm{vec}(T^{(k+1)}) = (\mathbf{L}^{(k)\top}W^{-1}\mathbf{L}^{(k)})^{-1}\mathbf{L}^{(k)\top}W^{-1}\mathrm{vec}(D),$$
where $\mathbf{L}^{(k)}=P^{(k)}\otimes I_d$.
8. $D^{(k)}=T^{(k)}P^{(k)\top}$.
9. **until** $\|D^{(k)}-D^{(k-1)}\|_{\mathrm{F}}/\|D^{(k)}\|_{\mathrm{F}}<\varepsilon$.
10. Output: $\hat{D}=D^{(k)}$.

**Note 2** (Initial approximation). The covariance matrix $W$ can partially be taken into account in the computation of the initial approximation by using the GTLS method, thereby improving the convergence rate as shown in Section 4. However, the initial approximation can also be chosen to be a rank $p$ TLS approximation $\hat{D}_{\mathrm{tls}}$ of $D$, as proposed in [31,32]: compute a rank $p$ truncated SVD approximation $D^{(0)}=\hat{D}_{\mathrm{tls}}=U_pS_pV_p^\top$ of $D$, with the SVD of the matrix $D$ equal to $USV^\top$, $U_p$ the truncation of the matrix $U$ to $m\times p$, $S_p$ the truncation of $S$ to $p\times p$ and $V_p$ the truncation of $V$ to $N\times p$, $U_p^\top U_p=I$, $V_p^\top V_p=I$ and let $\hat{T}_{\mathrm{tls}}=U_pS_p$, $\hat{P}_{\mathrm{tls}}^\top=V_p^\top$, such that $\hat{D}_{\mathrm{tls}}=\hat{T}_{\mathrm{tls}}\hat{P}_{\mathrm{tls}}^\top$.

The alternating least squares algorithm monotonically decreases the cost function value, so that it is globally convergent. The convergence rate, however, is linear and depends on the distribution of the singular values of $D$ [36, IV.A]. With a pair of closely spaced singular values the convergence rate could be rather low.

### 3.3. Algorithm of Premoli–Rastello

The algorithm of Premoli–Rastello, see [19,20] is derived for the special case when

$$W = \mathrm{blkdiag}(W_1,\ldots,W_N), \quad \text{where } W_i\in\mathbb{R}^{m\times m}.$$

The following input/output representation of the model is used

$$\mathrm{rank}(\hat{D})\le p \iff \hat{X}\hat{B}=\hat{Y}, \quad \text{where } \hat{D}^\top=[\,\hat{X}\ \hat{Y}\,]$$

with $\mathrm{col\,dim}(\hat{X})=p$, $\mathrm{col\,dim}(\hat{Y})=d$, and $m=p+d$. The low rank approximation problem (5) can be solved partially by minimizing analytically with respect to $\hat{D}$. In this way the following equivalent unconstrained optimization problem is derived:

$$\hat{B}^* = \arg\min_{\hat{B}} f(\hat{B}), \quad (10)$$

where

$$f(\hat{B}) = \sum_{i=1}^N d_i^\top R^\top(RW_i^{-1}R^\top)^{-1}Rd_i, \quad R=[\,\hat{B}^\top\ -I\,] \quad (11)$$

Define the residual matrix

$$E(\hat{B})=(RD)^{\top}=X\hat{B}-Y, \qquad E^{\top}(\hat{B})=[\,e_1(\hat{B})\cdots e_N(\hat{B})\,]$$

and partition $d_i$ and $W_i$ as follows:

$$d_i=\begin{bmatrix}x_i\\ y_i\end{bmatrix}\begin{matrix}p\\ d\end{matrix}, \qquad W_i=\begin{matrix}\begin{matrix}p & d\end{matrix}\\ \begin{bmatrix}W_{x,i} & W_{xy,i}\\ W_{yx,i} & W_{y,i}\end{bmatrix}\end{matrix}\begin{matrix}p\\ d\end{matrix}$$

The first order optimality condition $f'(\hat{B})=0$ of (10) is

$$2\sum_{i=1}^{N}(x_i e_i^{\top}(\hat{B})\Gamma_i^{-1}(\hat{B})-(W_{x,i}\hat{B}-W_{xy,i})\Gamma_i^{-1}$$
$$(\hat{B})e_i(\hat{B})e_i^{\top}(\hat{B})\Gamma_i^{-1}(\hat{B}))=0$$

where

$$\Gamma_i(\hat{B})=RW_i^{-1}R^{\top}.$$

We aim to find a solution of $f'(\hat{B})=0$ that corresponds to a solution of the low rank approximation problem, i.e., to a global minimum point of $f$.

The algorithm proposed in [19] uses an iterative procedure starting from an initial approximation $B^{(0)}$ and generating a sequence of approximations $B^{(k)}$, $k=0,1,2,\ldots$ that approaches a solution of $f'(\hat{B})=0$. The iteration is implicitly defined by the equation

$$F(B^{(k+1)},B^{(k)})=0, \tag{12}$$

where

$$F(B^{(k+1)},B^{(k)})=2\sum_{i=1}^{N}(x_i(B^{(k+1)\top}x_i-y_i)^{\top}\Gamma_i^{-1}(B^{(k)})$$
$$-(W_{x,i}B^{(k+1)}-W_{xy,i})\Gamma_i^{-1}(B^{(k)})e_i(B^{(k)})e_i^{\top}$$
$$\times(B^{(k)})\Gamma_i^{-1}(B^{(k)}))$$

Note that $F(B^{(k+1)},B^{(k)})$ is linear in $B^{(k+1)}$, so that $B^{(k+1)}$ can be computed in a closed form as a function of $B^{(k)}$. Eq. (12) with $B^{(k)}$ fixed can be viewed as a linear relaxation of the first order optimality condition of (10), which is a highly nonlinear equation.

An outline of the Premoli–Rastello algorithm is given in Algorithm 3. In general, solving the Eq. (12) for $B^{(k+1)}$ requires vectorization. The vec operator vectorizes a matrix column-wise and $\otimes$ denotes the Kronecker product. The following identity is used $\mathrm{vec}(XBY)=(Y^{\top}\otimes X)\mathrm{vec}(B)$ in order to transform (12) to the classical system of equations $G(B^{(k)})\mathrm{vec}(B^{(k+1)})=h(B^{(k)})$, where $G$ and $h$ are given in the algorithm. The mapping $b=\mathrm{vec}(B)\mapsto B$ is denoted by $\mathrm{vec}^{-1}$.

**Algorithm 3.** Algorithm of Premoli–Rastello.

1. Input: the data matrix $D\in\mathbb{R}^{m\times N}$, the covariance matrices $\{W_i\}_{i=1}^{N}$, a rank specification $p$, and a convergence tolerance $\varepsilon$.
2. Initial approximation: compute a GTLS solution $\hat{B}_{\mathrm{gtls}}$ of $XB\approx Y$ where $D^{\top}=[\,X\;Y\,]$ with $W_f=\frac{1}{N}\sum_{i=1}^{N}W_i$ where $W_i$ is the submatrix of $W$ at the intersection of the rows $(i-1)m+1$ to $i\cdot m$ and the columns $(i-1)m+1$ to $i\cdot m$ and let $B^{(0)}=\hat{B}_{\mathrm{gtls}}$ (see Algorithm 1).
3. Define: $D=\begin{bmatrix}x_1\cdots x_N\\ y_1\cdots y_N\end{bmatrix}\begin{matrix}p\\ d\end{matrix}$, where $d=m-p$.
4. $k=0$.
5. **repeat**
6. Let $G=0_{pd\times pd}$ and $h=0_{pd\times 1}$.
7. **for** $i=1,\ldots,N$ **do**
8. $e_i=B^{(k)\top}x_i-y_i$.
9. $N_i=\left(\begin{bmatrix}B^{(k)}\\ -I\end{bmatrix}^{\top}W_i\begin{bmatrix}B^{(k)}\\ -I\end{bmatrix}\right)^{-1}$.
10. $n_i=N_ie_i$.
11. $G=G+N_i\otimes(x_ix_i^{\top})-(n_in_i^{\top})\otimes W_{x,i}$.
12. $h=h+\mathrm{vec}(x_iy_i^{\top}N_i-W_{xy,i}n_in_i^{\top})$.
13. **end for**
14. Solve the system $Gb=h$ and let $B^{(k+1)}=\mathrm{vec}^{-1}(b)$.
15. $k=k+1$.
16. **until** $\|B^{(k)}-B^{(k-1)}\|_{\mathrm{F}}/\|B^{(k)}\|_{\mathrm{F}}<\varepsilon$
17. Output: $\hat{B}^{*}=B^{(k)}$.

**Note 3.** In the special cases of rank reduction by one, i.e., $p=m-1$, Eq. (12) becomes particularly simple. In this case, $\Gamma_i(b^{(k)})$, $e_i(b^{(k)})$, and $y_i$ are scalars, so that (12) can be written as

$$\sum_{i=1}^{N}(x_i(x_i^{\top}b^{(k+1)}-y_i)\Gamma_i^{-1}(b^{(k)})-(W_{x,i}b^{(k+1)}-W_{xy,i})$$
$$\Gamma_i^{-1}(b^{(k)})e_i(b^{(k)})e_i(b^{(k)})\Gamma_i^{-1}(b^{(k)}))=0,$$

which is equivalent to a standard linear system of equations $G(b^{(k)})b^{(k+1)}=h(b^{(k)})$,

$$\underbrace{\sum_{i=1}^{N}\left(\frac{x_ix_i^{\top}}{\Gamma_i(b^{(k)})}-W_{x,i}\frac{e_i^2(b^{(k)})}{\Gamma_i^2(b^{(k)})}\right)}_{G(b^{(k)})}b^{k+1}$$
$$=\underbrace{\sum_{i=1}^{N}\left(\frac{x_iy_i}{\Gamma_i(b^{(k)})}-W_{xy,i}\frac{e_i^2(b^{(k)})}{\Gamma_i^2(b^{(k)})}\right)}_{h(b^{(k)})}$$

**Note 4** (Relation to Gauss–Newton type algorithms). Algorithm 3 is not a Gauss –Newton type algorithm for solving the first order optimality condition because the approximation $F$ is not the first order truncated Taylor series of $f'$; it is a different linear approximation. The choice makes the

derivation of the algorithm simpler but complicates the convergence analysis.

**Note 5** (Convergence properties). Algorithm 3 is proven to be locally convergent with a super linear convergence rate, see [20, Section 5.3]. Moreover, the convergence rate tends to quadratic as the approximation gets closer to a minimum point. The algorithm, however, is not globally convergent and simulation results suggest that the region of convergence to a minimum point could be rather small. This requires a good initial approximation for convergence.

### *3.4. An algorithm based on classical local optimization methods*

Both the alternating least squares and the Premoli–Rastello algorithm are heuristic optimization methods. Next, we describe an algorithm for the low rank approximation problem based on classical local optimization methods. The classical local optimization methods have reached by now a high level of maturity. In particular, their convergence properties are well understood, while the convergence properties of the alternative methods are still not.

In order to apply classical optimization algorithms for the solution of (5), first we have to choose a parameterization of the model. A possible parameterization is given by the input/output representation, so the considered problem is (10).

A quasi-Newton type method requires an evaluation of the cost function $f(\hat{B})$ and its first derivative $f'(\hat{B})$. Both the cost function and its first derivative are available in closed form, so that their evaluation is a matter of numerical implementation of the involved operations. The computational steps are summarized in Algorithm 4. The proposed algorithm, based on a classical optimization method, is outlined in Algorithm 5.

**Algorithm 4.** Cost function and first derivative evaluation.

1. Input: $D \in \mathbb{R}^{m\times N}$, $\{W_i\}_{i=1}^N$, $p$, and $\hat{B}$.
2. Define: $D = \begin{bmatrix} x_1 & \cdots & x_N \\ y_1 & \cdots & y_N \end{bmatrix} \begin{matrix} p \\ d \end{matrix}$, where $d = m - p$.
3. Let $f = 0_{1\times 1}$ and $f' = 0_{p\times d}$.
4. **for** $i = 1, \ldots, N$ **do**
5. $e_i = \hat{B}^\top x_i - y_i$.
6. Solve the system $\left( \begin{bmatrix} \hat{B} \\ -I \end{bmatrix}^\top W_I \begin{bmatrix} \hat{B} \\ -I \end{bmatrix} \right) n_i = e_i$.
7. $f = f + e_i^\top n_i$.
8. $f' = f' + x_i n_i^\top - (W_{x,i}\hat{B} - W_{xy,i}) n_i n_i^\top$.
9. **end for**
10. Output: the cost function $f$ and its first derivative $2f'$ at $\hat{B}$.

**Algorithm 5.** Algorithm based on classical local optimization.

1. Input: the data matrix $D \in \mathbb{R}^{m\times N}$, the covariance matrices $W_i$, $i = 1, \ldots, N$, a rank specification $p$, and a convergence tolerance $\varepsilon$.
2. Initial approximation: compute a GTLS approximation $\hat{B}_{\text{gtls}}$ of $D$, and let $B^{(0)} = \hat{B}_{\text{gtls}}$. (See Step 2 of Algorithm 3.)
3. Apply a standard optimization algorithm, e.g., the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) quasi-Newton method, for the minimization of $f$ over $\hat{B}$ with initial approximation $B^{(0)}$ and with cost function and first derivative evaluation performed via Algorithm 4. Let $\hat{B}^*$ be the approximation found by the optimization algorithm upon convergence.
4. Output: $\hat{B}^*$.

The optimization problem (10) is a nonlinear least squares problem, i.e.,

$$f(\hat{B}) = F^\top(\hat{B})F(\hat{B})$$

for certain $F : \mathbb{R}^{p\times d} \to \mathbb{R}^{Nd}$. Therefore, the use of special optimization methods like the Levenberg–Marquardt method is preferable. The vector $F(\hat{B})$, however, is computed numerically, so that the Jacobian $J(\hat{B}) = \delta F_i/\delta \hat{B}_j$, where $\hat{B} = \text{vec}(\hat{B})$, cannot be found in closed form. A possible workaround for this problem is proposed in [37], where an approximation, called quasi-Jacobian is used instead. The quasi-Jacobian can be evaluated in a similar way to the one for the gradient, which allows use of the Levenberg–Marquardt method for the solution of the low rank approximation problem.

**Note 6** (Row versus column error covariances). Algorithms 3 and 5 are designed for the case when $m \geq N$ and row-wise correlated measurement errors. In chemometrics the data matrix usually has size $m \times N$ with $m \leq N$. When the measurement errors are uncorrelated or column-wise correlated, Algorithms 3 and 5 can be applied to the transposed data matrix. For other cases of measurement error correlation, the algorithms need to be optimized by considering the left kernel of $\hat{D}$, i.e., the following modifications of $\mathcal{C}2$ and $\mathcal{C}3$ should be used:

$\mathcal{C}2'$ : $\mathcal{N}^\top \hat{D} = 0$, where $\mathcal{N} \in \mathbb{R}^{m\times(m-p)}$ and $\mathcal{N}^\top\mathcal{N} = I$;
$\mathcal{C}3'$ : $[\, B^\top \; -I_{m-p} \,]\hat{D} = 0$, where $B \in \mathbb{R}^{p\times(m-p)}$.

## 4. Performance comparison of MLPCA versus EW-TLS

In this section the different approaches of the MLPCA and the EW-TLS method are compared with each other in order to discuss their individual weak and strong points. For the discussion, several simulated and experimental data sets from chemical measurements are used. The presented results are obtained by implementing the different algorithms in Mat-

lab (version 6.1) on a PC i686 with 800 MHz and 256 MB memory.

The data sets which are used differ in the structure of the covariance matrix $W$ of the measurement noise. In the first subsection we use data sets with uncorrelated measurement errors. The second subsection contains a data set with row-wise correlated measurement errors.

### 4.1. No correlation among the measurement errors

In this section we use data sets without any correlation among their measurement errors. For these data sets the weighting matrix $W$ in problem (5) is diagonal.

#### 4.1.1. Equal row error variances

In order to give results for the case of uncorrelated measurement noise with a diagonal covariance matrix $W_f$ which is equal for all rows, we use an experimental data set which was used previously by Wentzell and Andrews [31]. The reader is referred to the original work for a complete description of the experiment. In here, we only describe shortly the experiment.

**Example 1.** Thirty-one three-component mixtures containing toluene, chlorobenzene and heptane were derived from an augmented, three-level, three-factor, full factorial design. The spectra were obtained over the range 400–2500 nm on an NIRSystems model 6500 grating spectrometer at intervals of 2 nm and were the average result of 32 scans. Only standard deviations calculated from the replicate data for the first sample were available. The standard deviations for the first sample were used for all samples. So, the error covariance matrix was the same for every row of the measurement error matrix. Because there was no information about the correlations among the data, we used a diagonal covariance matrix for each row.

Because the error covariance matrix was the same for each row, the GTLS approach could be used. We applied both, the MLPCA and the GTLS, approaches to the noisy data matrix $D$ of size $31 \times 1050$ in order to estimate the error-free measurement data matrix $\hat{D}$ of rank 3. The MLPCA algorithm that we used for this data set is the standard MLPCA algorithm for uncorrelated errors (see [31, Table 1]). The GTLS algorithm that we applied is Algorithm 1 from the previous section. For both, the MLPCA and the GTLS, algorithms the relative error $\frac{\|D-\hat{D}\|_F}{\|D\|_F}$ and the computation time is presented in Table 2. The table demonstrates that GTLS is a practical alternative to MLPCA for this specific class of structure in the measurement error matrix. The measurement errors are uncorrelated and the covariance matrices of the errors in the rows are assumed to be equal for all rows. This class of measurement noise structure (uncorrelated errors with equal row error variance) is quite common for real-life data.

Table 2
MLPCA and GTLS for near-infrared spectroscopic data described in Example 1

| | Relative error | Time (s) |
|---|---|---|
| MLPCA | 0.15834292416862 | 37.6150 |
| GTLS | 0.15834292417331 | 4.1760 |

#### 4.1.2. Unequal row error variances

To discuss the performance of the MLPCA and the EW-TLS approach for data with uncorrelated measurement errors, we use simulated data sets.

**Example 2.** The data sets all have 10 rows, but the number of columns $N$ increases from 10 till 200 in steps of 10. So, this simulation set contains data sets of size $10 \times N$ and pseudorank 2. For each data set the error-free data matrix $D_0$ is generated by multiplying a $10 \times 2$ random matrix of full rank (with Matlab's `rand`) by a $2 \times N$ matrix that also is created with `rand`. The matrix of the measurement standard deviations corresponding to this $10 \times N$ matrix is determined by generating a $10 \times N$ matrix of uniform distributed random numbers between 0 and 0.01. This ensures that there is no pattern in the standard deviation matrix. Now, the $10 \times N$ matrix of measurement errors $\Delta D$ is generated by taking a $10 \times N$ random matrix with normally distributed elements (with Matlab's `randn`) and multiplying this, element-wise, by the standard deviation matrix. Finally, the noisy data matrix $D$ is the sum of the error-free matrix $D_0$ and the noise matrix $\Delta D$.

We applied the MLPCA method to the given data matrix $D$ and the EW-TLS method to the transposed data matrix $D^\top$ (see Note 6) to approximate the 'best' rank 2 approximation matrix $\hat{D}$ and the results are presented in Table 3 and Fig. 1. For the '*mlpca*' approach (see [31, Table 1]) the truncated SVD solution (see Note 2) is used as initial value. For the '*mlpca(gtls)*' approach the GTLS approximation (see Algorithm 1) is used as initial value instead of the truncated SVD solution. Table 3 and Fig. 1 also contain the results of applying the EW-TLS approach with two different algorithms. For '*ewtls*' Algorithm 3 is used and for '*ewtls(opt)*' Algorithm 5 is applied to the transposed data. For both algorithms we have used the GTLS approximation as initial starting point. Table 3 contains the relative error $\frac{\|D_0-\hat{D}\|_F}{\|D_0\|_F}$ after applying the four algorithms to a data set of size $10 \times 200$. The results in the table are the average of the relative error over 100 repetitions for different noise realizations. From the table it is clear that the '*ewtls*' Algorithm 3 diverges. It is very sensitive

Table 3
MLPCA and EW-TLS for data with uncorrelated measurement error described in Example 2

| Approach | Relative error |
|---|---|
| mlpca | 0.00262998780684 |
| mlpca(gtls) | 0.00267695997595 |
| ewtls | 0.29211400246120 |
| ewtls(opt) | 0.00263335308541 |

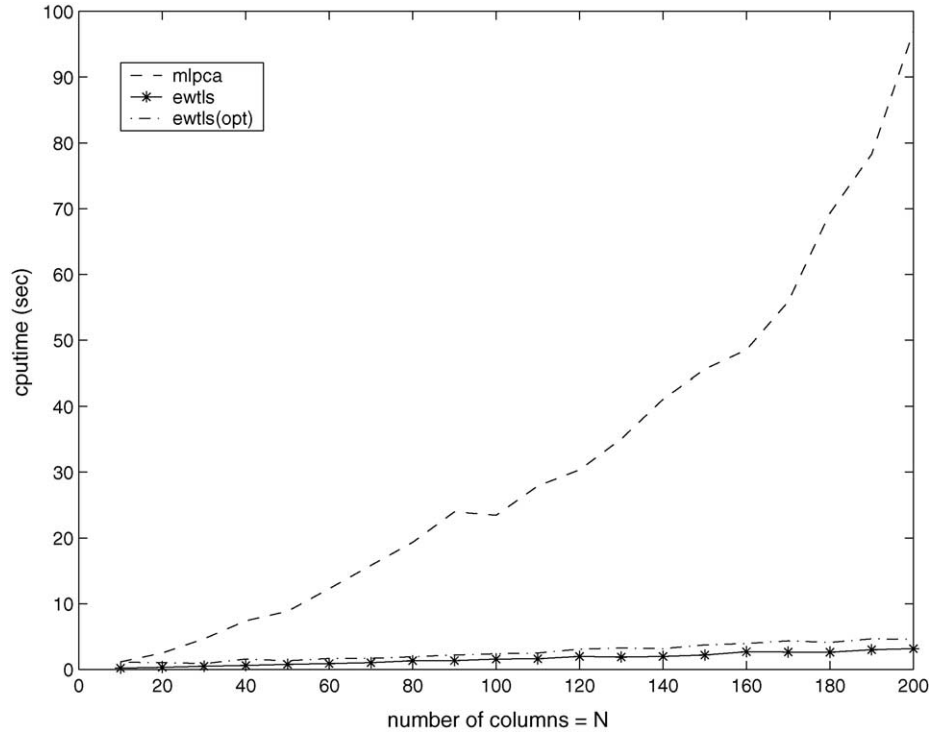


Fig. 1. The cputime of the MLPCA and the EW-TLS approaches for data with uncorrelated measurement error described in Example 2.

to local minima. Nevertheless, the '*ewtls(opt)*' Algorithm 5 converges to the same solution as the MLPCA algorithms do. Fig. 1 shows the computation time of the algorithms. For each $N$, the experiment is repeated 100 times for different noise realizations and the average execution times are reported. The EW-TLS method outperforms the MLPCA method for this specific case of measurement error. Especially for data sets of size $10 \times N$ with $10 \ll N$, the EW-TLS is much more efficient.

### *4.2. Correlation among the measurement errors along the rows only*

In calibration problems [32], the rows of the data matrix are formed by individual spectra. Hence, it is reasonable to discuss the case where correlations between the measurement errors exist only along the rows. So, the errors are uncorrelated along the columns of the data matrix. In these cases the full covariance matrix $W$ of the matrix of measurement errors will be block diagonal and can be stored as a sparse matrix. To calculate the inverse of $W$ the diagonal blocks can be inverted individually. For the discussion of this specific class of measurement structure a simulated data set from chemical measurements is used which was previously described in a paper by Wentzell and Lohnes [32].

**Example 3.** The simulated data set contained spectra from 10 samples of three-component mixtures. The concentration of each component in each of the 10 mixtures had a value between 0 and 1 from a uniform random number distribution. The spectral profiles of the three components were Gaussian with a standard deviation of 20 nm and maximum molar absorptivities at 480, 500 and 520 nm, respectively. Pure spectral vectors were generated between 400 and 600 nm at 20 nm intervals. The noise-free data matrix $D_0$ was calculated by multiplying the $10 \times 3$ matrix of concentrations by the $3 \times 11$ matrix of pure component spectra. To add a noise matrix $\Delta D$ of correlated measurement errors to the noise-free matrix $D_0$, first a $10 \times 11$ matrix $D'$ of uncorrelated measurement errors was generated. The matrix of the measurement standard deviations corresponding to this $10 \times 11$ matrix is determined by generating a $10 \times 11$ matrix of uniform distributed random numbers between 0 and 0.01. This ensures that there is no pattern in the standard deviation matrix. Now, the $10 \times 11$ matrix of uncorrelated measurement errors $D'$ is generated by taking a $10 \times 11$ random matrix with normally distributed elements (with Matlab's `randn`) and multiplying this, element-wise, by the standard deviation matrix. To introduce correlations among the errors within the rows, the rows of matrix $D'$ were filtered using a $1 \times 5$ moving average digital filter (see [31, Eqs. (34)–(36)] for the definition) to construct the correlated error matrix $\Delta D$. This error matrix $\Delta D$ w as added to the noise-free part $D_0$ in order to complete the noisy data matrix $D = D_0 + \Delta D$ of size $10 \times 11$.

To this simulated data set we have applied the MLPCA approach and the EW-TLS algorithms. Although the EW-

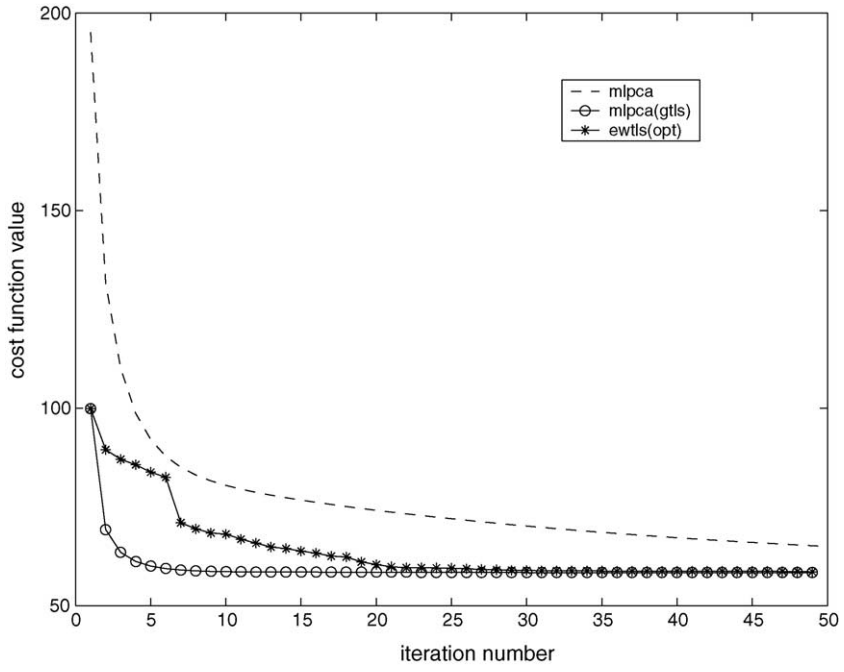


Fig. 2. The cost function value per iteration number is displayed for the MLPCA and the EW-TLS approaches applied to data with row error covariance only as described in Example 3.

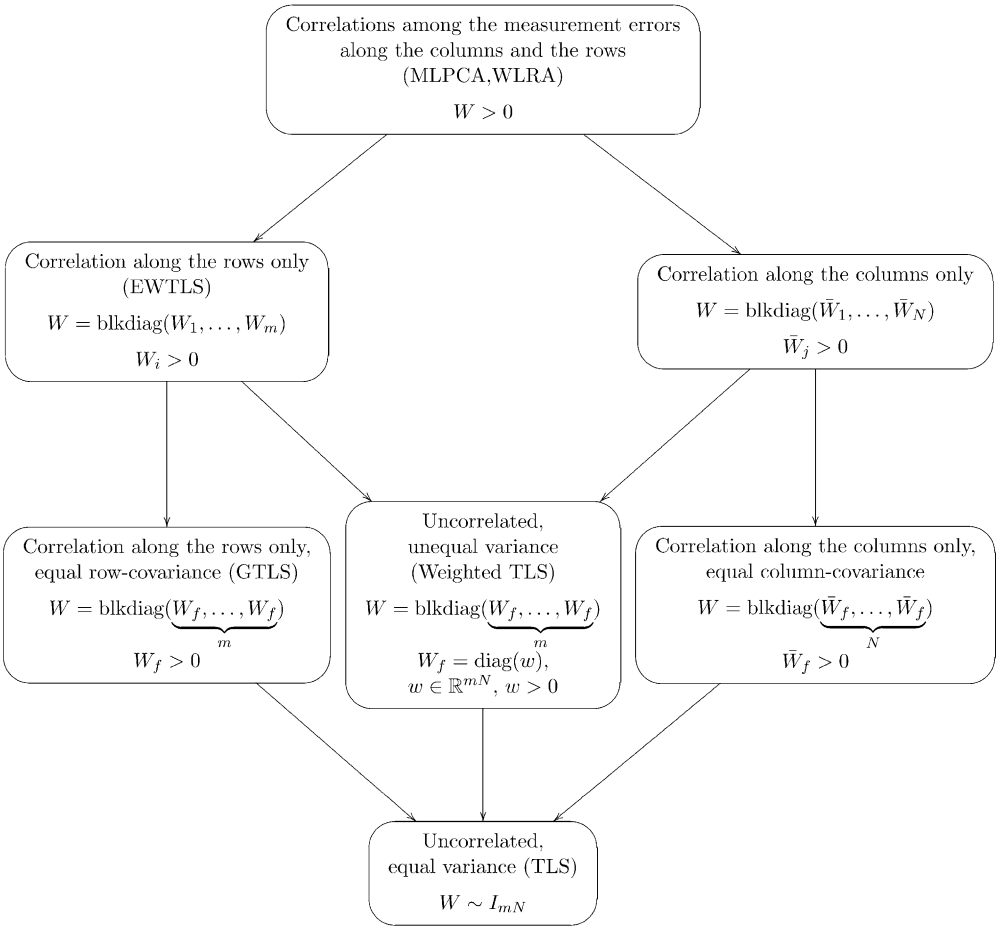


Fig. 3. Hierarchy of covariance structures of the measurement noise. Between brackets are given the problems that use the assumption.

Table 4
MLPCA and EW-TLS for data with row error covariance only as described in Example 3

| Approach | Cost ($\sim 10^5$) | Relative error |
|---|---|---|
| mlpca | 0.00058253884923 | 0.00472272972683 |
| mlpca(gtls) | 0.00058233611546 | 0.00468049099448 |
| ewtls Algorithm 3 | 9.21164965017561 | 0.44610730154588 |
| ewtls(opt) Algorithm 5 | 0.00058356157547 | 0.00478011794589 |

TLS algorithms are not yet optimized for handling correlated measurement errors in data sets where $m \ll N$, as is the case here (see Note 6), we nevertheless applied it to this data set. After all, from the previous section it is clear that the EW-TLS approach can be very useful for chemometrical problems. We compare the cost which is defined by $\mathrm{vec}^\top(D - \hat{D})W^{-1}\mathrm{vec}(D - \hat{D})$ and the relative error $\frac{\|D_0 - \hat{D}\|_F}{\|D_0\|_F}$. The results are presented in Table 4. The table contains the results of applying the MLPCA approach with two different initial starting points. For the '*mlpca*' approach the truncated SVD solution (see Note 2) is used. For the '*mlpca(gtls)*' approach the GTLS approximation (Algorithm 1) is used as initial value instead of the truncated SVD solution. We expect that the MLPCA algorithm with the GTLS approximation as initial value needs fewer iterations to converge. Fig. 2 shows this is indeed the case. Table 4 also contains the results of applying the EW-TLS approach with two different algorithms. These are the same as in the previous section. For '*ewtls*' Algorithm 3 is used and for '*ewtls(opt)*' Algorithm 5 is applied to the data. For both algorithms we have used the GTLS approximation as initial starting point. From the table it is obvious that we obtain the same solution as the MLPCA algorithms with the '*ewtls(opt)*' algorithm. Fig. 2 contains the results of the computation time of the MLPCA and the EW-TLS approaches. The '*mlpca(gtls)*' approach performs the best for this specific case of measurement error and matrix size. As said before, the EW-TLS method needs to be optimized for this kind of data set. Fig. 2 shows that it is better to use the GTLS approximation as initial value for the MLPCA algorithm.

## 5. Conclusions

In this paper we explored the tight equivalences between MLPCA and EW-TLS. We have shown that both methods can be used in any multivariate (ML) regression/calibration problem $XB \approx Y$ by reducing the problem to the same mathematical kernel problem, i.e.finding the closest (in a certain sense) weighted low rank matrix approximation where the weight is derived from the distribution of the errors in the data. Different solution approaches, as used in MLPCA and EW-TLS, have been discussed. From these discussions we can conclude that it is better to use the GTLS approximation as initial value for the MLPCA algorithm instead of the truncated SVD. Moreover, we have applied the EW-TLS algorithm to chemometrical data. The simulations show that the EW-TLS method certainly has potential for the problems in chemometrics. EW-TLS outperforms MLPCA for data sets with unequal row error variances. Nevertheless, a strong point of the MLPCA approach is its good convergence. But, the current algorithm is computationally not very efficient. Future work will focus on improving the computational efficiency of MLPCA by combining its good convergence with the efficiency of the EW-TLS algorithm.

## Acknowledgements

Prof. Dr. Sabine Van Huffel is a full professor, Mieke Schuermans and Ivan Markovsky are research assistants at the Katholieke Universiteit Leuven, Belgium. Prof. Dr. Peter D. Wentzell is a professor at the Dalhousie University, Canada, sponsored by the NSERC Canada and the Dow Chemical Company. Our research is supported by: Research Council KUL: GOA-Mefisto 666, IDO /99/003 and /02/009 (Predictive computer models for medical classification problemsusing patient data and expert knowledge), several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0078.01 (structured matrices), G.0407.02 (support vector machines), G.0269.02 (magnetic resonance spectroscopic imaging), G.0270.02 (nonlinear Lp approximation), research communities (ICCoS, ANMMM); IWT: PhD Grants; Belgian Federal Government: IUAP V-22 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling; EU: PDT-COIL, BIOPATTERN, ETUMOUR;

## References

[1] R.J. Adcock, A problem in least squares, The Analyst 4 (1877) 183–184.
[2] G.H. Golub, C.F. Van Loan, An analysis of the total least squares problem, SIAM J. Numer. Anal. 17 (1980) 883–893.
[3] S. Van Huffel, J. Vandewalle, The Total Least Squares Problem: Computational Aspects and Analysis, SIAM, Philadelphia, 1991.
[4] S. Van Huffel (Ed.), Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling, SIAM Proceedings Series, SIAM, Philadelphia, 1997.
[5] S. Van Huffel, P. Lemmerling (Eds.), Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications, Kluwer Academic Publishers, Dordrecht, 2002.
[6] G.H. Golub, Some modified matrix eigenvalue problems, Siam Rev. 15 (1973) 318–344.
[7] G.H. Golub, C.F. Van Loan, Matrix Computations, 3rd ed., The Johns Hopkins University Press, Baltimore, 1996.
[8] C.-L. Cheng, J.W. Van Ness, Statistical Regression With Measurement Error, Arnold, London, 1999.
[9] W.A. Fuller, Error Measurement Models, Wiley, New York, 1987.
[10] L.J. Gleser, Estimation in a multivariate "errors in variables" regression model: large sample results, Ann. Statist. 9 (1981) 24–44.
[11] R.D. Degroat, E.M. Dowling, The data least squares problem and channel equalization, IEEE Trans. Sign. Process. 41 (1993) 407–411.
[12] S. Van Huffel, J. Vandewalle, 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. Appl. 10 (3) (1989) 294–315.

[13] R.D. Fierro, G.H. Golub, P.C. Hansen, D.P. O'Leary, Regularization by truncated total least squares, SIAM J. Sci. Comp. 18 (1997) 1223–1241.
[14] D. Sima, S. Van Huffel, G.H. Golub, Regularized total least squares based on quadratic eigenvalue problem solvers, BIT (2004), in press.
[15] K.S. Arun, A unitarily constrained total least-squares problem in signal-processing, SIAM J. Matrix Anal. Appl. 13 (1992) 729–745.
[16] T.J. Abatzoglou, J.M. Mendel, G.A. Harada, The constrained total least squares technique and its applications to harmonic superresolution, IEEE Trans. Acoust. Speech Sign. Process. 39 (1991) 1070–1087.
[17] A. Kukush, I. Markovsky, S. Van Huffel, Consistency of the structured total least squares estimator in a multivariate model, J. Statist. Plann. Inference (2004), in press.
[18] N. Mastronardi, P. Lemmerling, S. Van Huffel, Fast regularized structured total least squares algorithm for solving the basic deconvolution problem, Numer. Lin. Alg. Appl. (2002), in press.
[19] A. Premoli, M.L. Rastello, The parametric quadratic form method for solving problems with element-wise weighting, in: S. Van Huffel, P. Lemmerling (Eds.), Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications, Kluwer, 2002, pp. 67–76.
[20] I. Markovsky, M.L. Rastello, A. Premoli, A. Kukush, S. Van Huffel, The element-wise weighted total least squares problem, Comp. Stat. Data Anal., in press. ftp://ftp.esat.kuleuven.ac.be/pub/SISTA/markovsky/abstracts/02–48.html.
[21] A. Kukush, S. Van Huffel, Consistency of elementwise-weighted total least squares estimator in a multivariate errors-in-variables model $AX = B$, Metrika 59 (1) (2004) 75–97.
[22] S. Wold, A. Ruhe, H. Wold, W.J. Dunn, The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses, SIAM J. Sci. Stat. Comput. 5 (3) (1984) 735–743.
[23] A. Burnham, R. Viveros, J. MacGregor, Frameworks for latent variable multivariate regression, J. Chemomet. 10 (1996) 31–46.
[24] S. de Jong, PLS fits closer than PCR, J. Chemomet. 7 (1993) 551–557.
[25] S. de Jong, Regression coefficients in multilinear PLS, J. Chemomet. 12 (1) (1998) 77–81.
[26] A. Phatak, Evalauation of some multivariate methods and their applications in chemical engineering, Ph.D. Thesis, University of Waterloo, 1993.
[27] A. Phatak, S. de Jong, The geometry of partial least squares, J. Chemomet. 11 (4) (1997) 311–338.
[28] P.D. Wentzell, D.T. Andrews, B.R. Kowalski, Maximum likelihood multivariate calibration, Anal. Chem. 69 (1997) 2299–2311.
[29] H. Martens, T. Naes, Multivariate Calibration, Wiley, New York, 1989.
[30] E.V. Thomas, Errors-in-variables estimation in multivariate calibration, Technometrics 33 (1991) 405–413.
[31] P.D. Wentzell, D.T. Andrews, D.C. Hamilton, K. Faber, B.R. Kowalski, Maximum likelihood principal component analysis, J. Chemomet. 11 (1997) 339–366.
[32] P.D. Wentzell, M.T. Lohnes, Maximum likelihood principal component analysis with correlated measurement errors: theoretical and practical considerations, Chemomet. Intell. Lab. Syst. 45 (1999) 65–85.
[33] J.T. Webster, R.F. Gunst, R.L. Mason, Latent root regression analysis, Technometrics 16 (1974) 513–522.
[34] R.F. Gunst, J.T. Webster, R.L. Mason, A comparison of least squares and latent root regression estimators, Technometrics 18 (1976) 75–83.
[35] W. Li, J. Qin, Consistent dynamic PCA based on errors-in-variables subspace identification, J. Process Contr. 11 (2001) 661–678.
[36] J.H. Manton, R. Mahony, Y. Hua, The geometry of weighted low rank approximations, IEEE Trans. Sign. Process. 51 (2) (2003) 500–514.
[37] P. Guillaume, R. Pintelon, A Gauss–Newton-like optimization algorithm for "weighted" nonlinear least-squares problems, IEEE Trans. Sign. Process. 44 (9) (1996) 2222–2228.