Level choice in truncated total least squares

Diana M. Sima, Sabine Van Huffel

Electrical Engineering Department, Katholieke Universiteit Leuven, Belgium

Abstract

The method of truncated total least squares is an alternative to the classical truncated singular value decomposition used for the regularization of ill-conditioned linear systems. Truncation methods aim at limiting the contribution of noise or rounding errors by cutting off a certain number of terms in an expansion such as the singular value decomposition. To this end a truncation level k must be carefully chosen.

The truncated total least squares solution becomes more significantly dominated by noise or errors when the truncation level k is overestimated than the truncated singular value decomposition solution does. Model selection methods that are often applied in the context of the truncated singular value decomposition are modified in this paper to be applied in the context of the truncated total least squares. The proposed modified generalized cross validation combined with the truncated total least squares method performs better than the classical generalized cross validation combined with the truncated singular value decomposition, especially when both the coefficient matrix and the right-hand side are contaminated by noise.

Key words: truncated singular value decomposition, truncated total least squares, filter factors, effective number of parameters, model selection, generalized cross validation, information criteria

1 Introduction

Ill-posed problems are problems where the solution does not depend continuously on the input data, where arbitrarily small perturbations in the input data produce arbitrarily large changes in the solution. Discrete ill-posed problems might arise from the discretization of continuous problems, such as integral or differential models. In this paper we consider discrete linear ill-posed

Email addresses: diana.sima@esat.kuleuven.be (Diana M. Sima), sabine.vanhuffel@esat.kuleuven.be (Sabine Van Huffel).

problems, which are linear systems $Ax \approx b$, possibly incompatible, where the coefficient matrix A of dimension $m \times n$ ($m \geq n$) and the right-hand side vector b of length m are the given data, while the vector x is the solution to be computed.

Extensive characterizations of discrete ill-posed problems of the type $Ax \approx b$ exist in the literature (see the book (10) and the references therein). It is known, for instance, that the properties of such a system can be understood from the singular value decomposition (7) of A. In an ill-posed system there is no clear gap in the decay of the singular values of A. This makes it difficult to decide the best value that can be used as numerical rank. Moreover, it usually happens that the singular vectors corresponding to smaller and smaller singular values have increasing complexity (meaning that they contain more and more sign changes, oscillations). If the least squares or total least squares solution of $Ax \approx b$ is computed, then these oscillations are emphasized in the solution by the noise in the data or by the effects of working in finite precision arithmetic. Therefore, ill-posed problems cannot be solved directly; some form of *regularization*, which changes the original problem formulation, is needed.

In this paper we revisit one of the simplest families of regularization methods for discrete linear ill-posed problems: truncation methods. We shall focus on the *truncated singular value decomposition* (TSVD) and the *truncated total least squares* (TTLS) methods. These methods depend on a parameter called the *truncation level*; this single parameter has an important influence on the computed solutions of the truncation methods.

A typical assumption of the existing methods for choosing the truncation level (or, more generally, a regularization parameter for the regularization of the discrete linear ill-posed problem $Ax \approx b$) is that b is the sum of a b^{exact} and a b^{noise}, where b^{exact} is in such a way that $Ax = b^{\text{exact}}$ is compatible (and has as solution the true solution x^{exact}), and b^{noise} is a vector with zero mean i.i.d. random elements. Well-known methods such as the discrepancy principle, cross validation and generalized cross validation, information criteria, the L-curve (see (10) for an overview), and also newer analyses such as the residual exploitation technique in (11), belong to this category of methods which assume that the matrix A is exactly known.

The aim of our study is to extend several classical methods for model selection to the case when the matrix A is noisy as well. The specific conditions under which the methods here will be most favorable are the errors-in-variables assumptions that both the matrix A and the right-hand side vector b are contaminated by uncorrelated i.i.d. zero mean errors of equal variance. If this is not the case, then the data should be scaled *a priori*. Scaling recommendations are given in (17, §3.6.2, pp.90-91).

In our analysis we complement the discussion in (4)—the paper that first proposed and analyzed the truncated total least squares as a regularization method; in (4) the methods used for choosing the truncation level are either the L-curve or the minimization of $\|Ax - b\|$. We propose modifications of other methods such as the generalized cross validation criterion and Akaike's information criterion. We shall make extensive use of the *filter factors* formulation of the truncated total least squares solution proved in (4). Intuitively explained, the idea of using the special filter factors corresponding to TTLS is that these positive numbers contain much more information about the solution than the filter factors in TSVD do (since the latter have only two possible values: 0 and 1).

The property that we implicitly use in the modified criteria for choosing the truncation level is that the TTLS solution is more sensitive to the noise in the data or to rounding errors than the TSVD solution is, when the truncation level is greater than a certain critical value. We show how this property is linked to the values of the TTLS filter factors. A surprising behavior illustrated by our simulations is that the criteria for TTLS sometimes work better than the classical criteria for TSVD even if the errors-in-variables assumption is not satisfied (*e.g.*, if only b is noisy or if the problem is noiseless). We explain this again by the better ability of the TTLS criteria to monitor the higher sensitivity of the solution to rounding errors.

In the following section we present the background on the two truncation algorithms under consideration: truncated singular value decomposition and truncated total least squares. These methods are used as regularization methods in the context of discrete linear ill-posed problems; for this reason, we give in §2.2 an illustration of such an ill-posed problem and we show the usefulness of the truncation methods. The issue of choosing the truncation level is discussed in section 3, where we first review the most commonly used approaches with the truncated singular value decomposition. Then we analyze several proposed modifications for truncated total least squares. Section 4 confirms the effectiveness of one of the most relevant proposed approaches (modified generalized cross validation) on a benchmark of simulation examples.

As usual, we denote vectors with lowercase letters and matrices with capital letters, the vector 2-norm by $\|v\|^2 = v^\top v$, while for matrices $\|\cdot\|_F$ denotes the Frobenius norm; we use the superscript $\top$ for the transpose and the superscript $\dagger$ for the pseudoinverse of a matrix.

2.1 Truncation methods for linear estimation

Truncation methods are an effective way of eliminating unwanted subspaces (such as the noise subspace) from given data sets, using the singular value decomposition (SVD) (7) or other related factorizations. For instance, if for a certain application the noise level is *a priori* known, then it is reasonable to eliminate all information that is below the noise level. It may happen that useful information from the problem is also embedded in that noise; this is an intuitive explanation why *truncation methods* might be too drastic, and thus they are not always the most appropriate methods for all problem settings.

Linear approximation problems of the form $Ax \approx b$ are classically solved either in the least squares (LS) sense by minimizing the changes on b that make the system compatible, or in the total least squares sense (TLS (6; 17)) by minimizing the changes on A and b that make the system compatible. These approaches generally do not work well when the system is ill-posed, as will be illustrated with an example in §2.2. Modified estimation methods based on truncation for the least squares and total least squares settings are, respectively, the truncated singular value decomposition (see Algorithm 1) and the truncated total least squares (see Algorithm 2).

Algorithm 1 Truncated SVD (TSVD).

1: Compute the SVD of $A = U'\Sigma'{V'}^\top$, where U' is $m \times n$, V' is $n \times n$ and ${U'}^\top U' = {V'}^\top V' = V'{V'}^\top = I_n$, and Σ' is an $n \times n$ diagonal matrix with the singular values $\sigma'_1 \geq \sigma'_2 \geq \ldots \geq \sigma'_n \geq 0$ on the diagonal;
2: Choose an appropriate truncation level $k \leq n$;
3: Compute the truncated SVD solution

$$x_{\mathrm{TSVD},k} = V'{\Sigma'_k}^\dagger {U'}^\top b, \qquad \text{with } \Sigma'_k = \mathrm{diag}\{\sigma'_1, \ldots, \sigma'_k, \underbrace{0, \ldots, 0}_{n-k}\}. \tag{1}$$

For truncating in the least squares context, the TSVD Algorithm 1 cuts off the contribution of the $n - k$ singular triplets of A corresponding to the smallest $n - k$ singular values of A from the least squares solution of $Ax \approx b$.

In the total least squares context the TTLS Algorithm 2 finds the minimum 2-norm solution x of $A_k x = b_k$, where $[A_k\ b_k]$ is the best rank-k approximation of $[A\ b]$. The formula (2) of the TTLS solution follows from the fact that all

Algorithm 2 Truncated TLS (TTLS).

1: Compute the SVD of $[A\ b] = U\Sigma V^\top$, where U is $m \times (n+1)$, V is $(n+1)\times(n+1)$ and $U^\top U = V^\top V = VV^\top = I_{n+1}$, and Σ is an $(n+1)\times(n+1)$ diagonal matrix with the singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq \sigma_{n+1} \geq 0$ on the diagonal;
2: Choose an appropriate truncation level $k \leq n$.
3: Let $[A_k\ b_k] := U\Sigma_k V^\top$, where Σ_k is Σ with the smallest $n-k+1$ singular values set to zero, $\Sigma_k = \mathrm{diag}\{\sigma_1, \ldots, \sigma_k, 0, \ldots, 0\}$;
4: Solve for the minimum 2-norm solution of the truncated problem $A_k x \approx b_k$ and obtain the truncated TLS solution

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

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

$$V = \begin{array}{c} \overset{k}{\longleftrightarrow}\ \overset{l}{\longleftrightarrow} \\ \begin{bmatrix} V_{11}^k & V_{12}^k \\ V_{21}^k & V_{22}^k \end{bmatrix} \begin{array}{l} \updownarrow n \\ \updownarrow 1 \end{array} \end{array} \tag{3}$$

solutions of $A_k x = b_k$ should satisfy

$$\begin{bmatrix} x \\ -1 \end{bmatrix} \in \mathrm{Range} \begin{bmatrix} V_{12}^k \\ V_{22}^k \end{bmatrix},$$

and thus any solution x should have the form

$$x = -\frac{V_{12}^k}{\|V_{22}^k\|^2} \left((V_{22}^k)^\top + w \right),$$

where the vector w is any vector orthogonal to $(V_{22}^k)^\top$. The minimum norm solution is obtained for $w = 0$.

In Algorithm 2 there is a necessary condition that the vector V_{22}^k is not all zero. V_{22}^k is zero in *nongeneric* cases (see (17, §3.4.1)), but it can be nearly zero in close-to-nongeneric cases, which can happen in ill-posed problems. As noted in (4), we can always decrease k enough such that (the longer) V_{22}^k has large enough norm. Therefore, if V_{22}^k is nearly zero (for some k), this will be an indication that we are using a too big truncation level k.

For later use, we recall the interlacing property for singular values (7), which guarantees the following ordering of the singular values of A and $[A\ b]$:

$$\sigma_1 \geq \sigma_1' \geq \sigma_2 \geq \sigma_2' \geq \cdots \geq \sigma_n \geq \sigma_n' \geq \sigma_{n+1}. \tag{4}$$

Before going further we illustrate with an example of an ill-posed problem the fact that it is not advisable to use the classical estimation methods (least squares, total least squares) directly.

A matrix A_0 and a vector b_0 with real valued elements and dimensions $m = 200$, $n = 30$, are simulated using the example `ilaplace` from the Regularization Tools (9). This example that originally comes from (18) constructs A_0 as a discretization of the inverse Laplace transform, using Gauss-Laguerre quadrature for the discretization of the inverse Laplace integral operator. In (18) and in Hansen's toolbox (9), the sampling points where the integral equation is evaluated were as many as the quadrature points, thus yielding a square discretized system. We slightly modify this setting in order to construct rectangular systems, by allowing more sampling points than quadrature points.

In this simulation, $A_0 x \approx b_0$ has a certain exact solution vector x_0, satisfying the equation $A_0 x_0 = b_0$. We construct an incompatible noisy example $Ax \approx b$, by adding Gaussian white noise with a standard deviation of 1e-5 to all elements of A_0 and b_0. Figure 1 shows the singular values of A_0 as they decay without a noticeable gap towards machine precision. For comparison, Figure 1

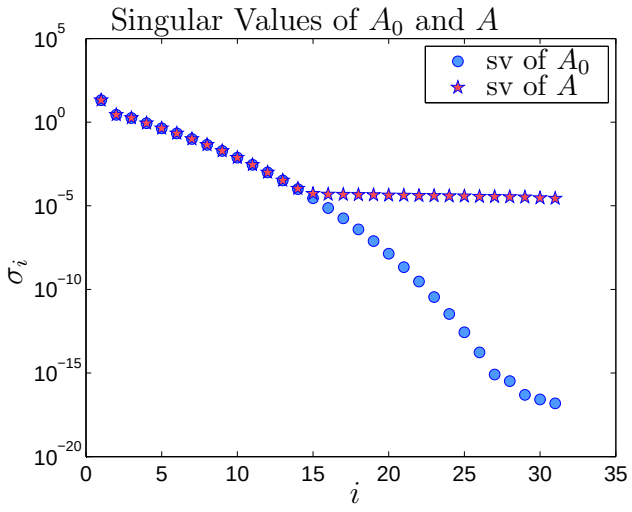


Fig. 1. Singular values of A_0 decaying without gap towards zero, where the data A_0 comes from a discrete ill-posed problem. Singular values of the noisy data A level off at the noise level.

also shows the decaying singular values of the noisy A and it is noticeable that the smallest singular values stagnate at a certain level, determined by the standard deviation of the added noise.

The conditioning of the matrix A (about 2e+5) is much better than that of A_0 (about 3e+17), but this is a misleading improvement. In fact, the added

noise has a disastrous effect on the problem of estimating x using, *e.g.*, the LS or TLS method.

Figure 2 shows the true solution x_0 together with the LS and TLS solutions computed from the data $[A\ b]$. Note that both estimated solutions are very noisy and do not reconstruct the true solution well. The TLS solution is noticeably more contaminated than the LS solution.

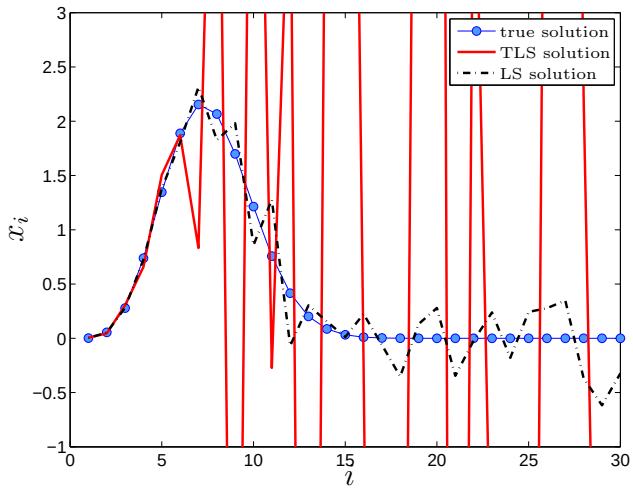


Fig. 2. The true solution x_0 is chosen to be the discretization of a smooth function. However, the LS and TLS solutions x computed from the noisy data $[A\ b]$ are influenced by the noise and are far from being reasonable approximations of the underlying true solution x_0.

Next we aim to illustrate that truncation is a good regularization method for the type of problem described before. We used the SVD of the noisy A and, respectively, $[A\ b]$ and we computed truncated SVD and truncated TLS solutions ($x_{\mathrm{TSVD},k}$ and $x_{\mathrm{TTLS},k}$) for all possible values of k, from 1 to $n = 30$. We plot in Figure 5 several reconstructed solutions, together with the exact x_0. Some facts are to be noted:

- the "non-truncated" solutions $x_{\mathrm{TSVD},n}$ and $x_{\mathrm{TTLS},n}$ are identical respectively to the LS and TLS solution in Figure 2. The truncated solutions with large k indices are also very much influenced by the noise and resemble these LS/TLS solutions;
- the truncated solutions with very small k level are too simplistic to capture the characteristics of the true solution x_0;
- there are several values of k, such as $k = 10$ in Figure 5, for which the truncated SVD and truncated TLS solutions are very good reconstructions of the true solution x_0.

Although we do not plot it, the noiseless problem $A_0x = b_0$ is also inaccurately solved numerically, unless we use truncation. The range of good truncation

levels is wider than in the noisy case, but the best reconstruction of x_0 is obtained for $k = 20$, for which singular values below 1e-10 are discarded from the SVD of A or $[A\ b]$.

The aim of regularization by truncation is thus to appropriately identify a good truncation level, and to construct a truncated solution that can capture the essential features of the unknown true solution, without explicit knowledge about the true solution, and even without *a priori* knowledge about the magnitude of the noise in the data.

3 Choice of the truncation level

This section is devoted to the main goal of the paper: choosing good truncation levels for truncated total least squares. We shall use several existing criteria for TSVD to inspire our choice of TTLS criteria.

We start in §3.1 with an overview of the most commonly used model selection criteria that are designed for the regularization of discrete linear ill-posed problems, and particularly used with TSVD. Then in §3.2 we discuss how each of the presented criteria can be modified to better suit the TTLS approach, and which other alternative methods are available in the literature to accommodate the case when A is noisy. We then focus on the generalized cross validation and Akaike's information criterion and show how their commonly used formulations can be simply transformed to a TTLS version.

Further on, in section 4, we illustrate that these latter criteria can give superior results compared to the original TSVD versions.

3.1 Classical criteria for TSVD

For the TSVD formulation, and assuming that only b is noisy, many classical model selection methods exist. Some basic ideas are common to all methods; for instance, there is a trade-off between the goodness of fit and the parsimony of the model. We now enumerate several well-known criteria. (For a more detailed discussion we refer to (10, Chapter 7).)

The discrepancy principle

This method that originates in (14) requires knowledge about the statistical properties of the noise that perturbs the data. Specifically, let us assume that the noise is i.i.d. with zero mean and variance σ^2. The norm of the residual

vector is computed for the candidate truncation level k and the following criterion is monitored:

$$\min\ k \quad \text{such that} \quad \|Ax_{\mathrm{TSVD},k} - b\|^2 \leq m\sigma^2, \tag{5}$$

thus, a truncation level is sought that provides a residual of the same magnitude as the noise in the data.

The L-curve

The norm of the truncated solution $\|x_{\mathrm{TSVD},k}\|$ is plotted against the norm of the residual $\|Ax_{\mathrm{TSVD},k} - b\|$ in log-log scale for various k's, and the *corner* in log-log scale is chosen (8). A variant is the *rotated L-curve* (16), where the product $\|Ax_{\mathrm{TSVD},k} - b\|\|x_{\mathrm{TSVD},k}\|$ is minimized for k. In log-log scale, the plot of this function looks like a rotated L-curve, such that the L-curve's corner becomes a minimum.

Generalized cross validation

Generalized cross validation (GCV) (3; 5; 19) is a method derived from the leave-one-out cross validation criterion. In the TSVD case it can be written as the following minimization problem:

$$\min_k \frac{\|Ax_{\mathrm{TSVD},k} - b\|^2}{(m-k)^2}. \tag{6}$$

Akaike's information criterion

For TSVD the classical Akaike's information criterion (AIC) (1) is defined as the minimization problem:

$$\min_k\ \log\left(\|Ax_{\mathrm{TSVD},k} - b\|\right) + \frac{k}{m}. \tag{7}$$

We notice an interesting relation between the function minimized by Akaike's information criterion and the GCV function. The logarithm of the GCV function is a multiple of

$$\log\left(\|Ax_{\mathrm{TSVD},k} - b\|\right) - \log\left(1 - \frac{k}{m}\right) - \log m.$$

Whenever k/m is small, the approximation $\log(1 - \frac{k}{m}) \approx -\frac{k}{m}$ holds. Thus, if the minima in (6) and (7) occur for small enough values of k, then the optimal solutions coincide.

In the TTLS case the consideration that A can be noisy implies several changes in the model selection criteria.

The norm of the residual in this case is defined as

$$\|R_k^{\mathrm{TLS}}\|_F^2 := \|[A\ b\,] - [A_k\ b_k\,]\|_F^2 \tag{8}$$

where $[A_k\ b_k]$ denotes the best rank-k approximation of $[A\ b\,]$.

The **discrepancy principle** could be generalized such that it uses this measure. If the elementwise noise in A and b is zero mean i.i.d. with variance σ^2, then the estimate of a 'denoised' rank-k approximation can be computed with the criterion:

$$\min\ k \quad \text{such that } \|R_k^{\mathrm{TLS}}\|_F^2 \leq m(n+1)\sigma^2. \tag{9}$$

The factor $m(n+1)$ takes care of the fact that there are $m(n+1)$ noisy elements in A and b. This simple generalization of the classical discrepancy principle does not share the statistical motivation of its counterpart in the LS case. In the ordinary regression setting when A is exact and only noise on b^{exact} is added, the LS residual preserves the (Gaussian) distribution of the assumed noise added to b^{exact}. In the TLS sense, the low rank residual matrix $[A\ b\,] - [A_k\ b_k\,]$ does not preserve the statistical properties of the added noise on the true A^{exact} and b^{exact} data that generates A and b.

In the **L-curve** criterion, we plot $\|R_k^{\mathrm{TLS}}\|_F$ (instead of $\|Ax_{\mathrm{TTLS},k} - b\|$) against $\|x_{\mathrm{TTLS},k}\|$ in log-log scale and choose the truncation level k corresponding to the corner.

In §3.4 we describe the adapted versions of the generalized cross validation and Akaike's information criteria, but for the next few paragraphs we would like to review several other approaches that are linked to TTLS through the fact that a low rank approximation of a noisy data matrix is sought. The truncation level plays the same role as the rank of the approximation matrix.

In signal processing there are many applications dealing with noisy signals. Many denoising methods proceed by stacking a signal in a (structured) matrix, which has the property that it is rank-deficient if the signal is noise-free, but it is full-rank for noisy signals. To relate to our previous notation, we view $[A\ b\,]$ as a single data matrix C consisting of a true mean matrix C^{true}, which is rank-deficient, plus a matrix C^{noise}, which makes $C^{\mathrm{true}} + C^{\mathrm{noise}}$ full rank. The (column) span of C^{true} is called the *signal subspace* and the (column) span of C^{noise} is called the *noise subspace*.

Denoising the signal boils down to a matrix decomposition followed by truncation. However, the dimension of the signal subspace is usually unknown beforehand.

A method of rank determination that aims at separating the signal subspace and the noise subspace is based on signal-to-noise ratio considerations. The signal-to-noise ratio of the best rank j approximation of $[A\ b]$ has the following definition (17, p. 89):

$$\mathrm{SNR}_j := \frac{\sigma_j([A_0\ b_0])^2}{m\sigma^2}, \tag{10}$$

where $\sigma_j([A_0\ b_0])$ is the jth singular value of the exact but unobservable $m \times (n+1)$ matrix $[A_0\ b_0]$, and σ^2 is the variance of the i.i.d. zero mean additive noise in $[A\ b]$. Then the optimal choice for the truncation level k such that the SVD components with signal-to-noise ratios smaller than 1 are discarded is (17, p. 89): choose k as the number of singular values σ_j that satisfy $\sigma_j > \sqrt{2m}\sigma$.

When the true noise variance σ^2 is unknown, the squared smallest singular value of $[A\ b]$ divided by the row dimension m is sometimes used as an estimate. Thus we can often successfully apply the criterion: choose k as the number of singular values σ_j that satisfy $\sigma_j > \sqrt{2}\sigma_{\min}([A\ b])$. (For the optimality of this choice see also (2).) The disentanglement between the signal subspace and the noise subspace will be accurate when there is a clear distinction between signal and noise (which translates into a clear gap in the singular values of the matrix).

Other choices of the signal subspace dimension used in practice are versions of Akaike's information criterion and of the minimum description length principle (20). We do not go into the details of these methods, since the purpose of all these methods is to find a good reconstruction of the signal subspace (equivalent to a lower rank reduction), but our interest in the context of the discrete ill-posed $Ax \approx b$ problem is primarily the estimation of a good solution x, without enforcing strong statistical assumptions and without developing an asymptotic method.

3.3 Filter factors

In the rest of Section 3 we aim at adapting the generalized cross validation criterion and Akaike's information criterion for TTLS. The basis of our proposed truncation level choice comes from the concept of *filter factors* of a regularized model. We employ an observation already anticipated in the introduction: the filter factors of TTLS are much more informative than the filter factors of TSVD.

In this subsection we denote by k a fixed truncation level in the TSVD or the TTLS algorithms.

The TSVD solution has a simple interpretation in terms of filter factors. If we set

$$f'_1 = f'_2 = \ldots = f'_k = 1, \quad \text{and} \quad f'_{k+1} = \ldots = f'_n = 0, \tag{11}$$

then the TSVD solution with truncation level k is simply, see (1):

$$x_{\mathrm{TSVD},k} = \sum_{i=1}^{n} f'_i \frac{{u'_i}^\top b}{\sigma'_i} v'_i, \tag{12}$$

where (with the notation from Algorithm 1), (u'_i, σ'_i, v'_i) for $i = 1, \ldots, n$ are the singular triplets of the matrix A. Thus, using only filter factors equal to one and zero, the TSVD solution is a filtered version of the ordinary least squares solution.

In general, a regularized solution to $Ax \approx b$ can be written as

$$x^{\mathrm{reg}} = \sum_{i=1}^{n} f_i \frac{{u'_i}^\top b}{\sigma'_i} v'_i, \tag{13}$$

with $f_i \in [0, 1]$. The filter factor formulation relates a regularized solution to the least squares solution of $Ax \approx b$. Moreover, the predicted right-hand side that corresponds to x^{reg} is:

$$b^{\mathrm{reg}} := Ax^{\mathrm{reg}} = U'\Sigma'{V'}^\top x^{\mathrm{reg}} = U'\mathrm{diag}\{f_1, \ldots, f_n\}{U'}^\top b. \tag{14}$$

In (4) it is proved that the TTLS solution $x_{\mathrm{TTLS},k}$ in (2) can be expressed in the filter factor formulation (14), using the filter factor values:

$$f_i = \sum_{\substack{j=k+1 \\ \sigma_j \neq \sigma'_i}}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{{\sigma'_i}^2}{{\sigma'_i}^2 - {\sigma_j}^2} \right), \qquad i = 1, \ldots, n, \tag{15}$$

where $v_{n+1,j}$ denotes the $(n+1, j)$ element of V, in other words, the last element of the right singular vector of $[A\ b]$ corresponding to the jth singular value. Here the symbols without the prime sign correspond to the SVD of $[A\ b]$ given in Algorithm 2. Note that using these filter factors we are able to express the TTLS solution with truncation level k in terms of the SVD of the matrix A. For the following theory it is useful to recall the interlacing property for singular values (4).

A theorem in (4) provides the following bounds on the TTLS filter factors:
Lemma 1 (Th. 4.2, (4))**.** *With the notation in Algorithms 1 and 2 assume $\sigma'_k \neq \sigma_{k+1}$. The first k filter factors f_i in (15) form a monotonically increasing*

sequence and satisfy

$$1 \le f_i \le 1 + \frac{\sigma_{k+1}^2}{\sigma_i'^2 - \sigma_{k+1}^2}, \qquad i = 1, \ldots, k \tag{16}$$

while the last $n-k$ filter factors satisfy

$$0 \le f_i \le \|V_{22}^k\|^{-2} \frac{\sigma_i'^2}{\sigma_k^2 - \sigma_i'^2}, \qquad i = k+1, \ldots, n. \tag{17}$$

These results show that $f_1, \ldots, f_k$ are greater than but close to 1, while $f_{k+1}, \ldots, f_n$ are small, provided that $\|V_{22}^k\|$ is not close to zero and that there are big enough gaps in the ordering $\sigma_i' > \sigma_{k+1}$ (for $i = 1, \ldots, k$) and $\sigma_k > \sigma_i'$ (for $i = k+1, \ldots, n$). In particular, f_k is close enough to one and f_{k+1} drops to a small value if there is a large gap in the middle of the following interlacing relationship:

$$\sigma_k \ge \sigma_k' \gg \sigma_{k+1} \ge \sigma_{k+1}'. \tag{18}$$

In this situation the TTLS solution resembles the TSVD solution with the same truncation level k.

However, whenever some of the filter factors for TTLS are greater than one, the TTLS solution is partly a "de-regularized" solution. We provide the following lemma containing opposite inequality relations, which show how large the filter factors may get if the above mentioned conditions on the magnitude of $\|V_{22}^k\|$ and the separation of certain singular values are not met.

Lemma 2. *With the notation in Algorithms 1 and 2 assume $\sigma_k' \ne \sigma_{k+1}$. Then the first k filter factors f_i in (15) form a monotonically increasing sequence and satisfy*

$$f_i \ge 1 + \frac{\sigma_{n+1}^2}{\sigma_i'^2 - \sigma_{n+1}^2}, \qquad i = 1, \ldots, k, \tag{19}$$

and the last $n-k$ filter factors form a monotonically decreasing sequence and satisfy

$$f_i \ge \frac{\|V_{21}^k\|^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_1^2 - \sigma_i'^2} \right), \qquad i = k+1, \ldots, n. \tag{20}$$

Proof. In relation (15) we use the ordering of the singular values $\sigma_j \ge \sigma_{n+1}$, for $j \le n$, together with the fact that $\sum_{k+1}^{n+1} v_{n+1,j}^2 = \|V_{22}^k\|^2$, and the knowledge that $v_{n+1,j} = 0$ whenever there is an i such that $\sigma_j = \sigma_i'$ (17, Th. 3.11). Then

$$\begin{aligned} f_i &= \sum_{j=k+1}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_j^2} \right) \ge \sum_{j=k+1}^{n+1} \frac{v_{n+1,j}^2}{\|V_{22}^k\|^2} \left(\frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_{n+1}^2} \right) = \frac{\sigma_i'^2}{\sigma_i'^2 - \sigma_{n+1}^2} \\ &= 1 + \frac{\sigma_{n+1}^2}{\sigma_i'^2 - \sigma_{n+1}^2}, \end{aligned}$$

proving (19). For the second inequality we use another characterization of the filter factors, given also in (4), namely,

$$f_i = \|V_{22}^k\|^{-2} \sum_{j=1}^{k} v_{n+1,j}^2 \left(\frac{\sigma_i'^2}{\sigma_j^2 - \sigma_i'^2} \right). \tag{21}$$

As opposed to (15) this formulation makes it clear for $i > k$ that, due to interlacing, all terms are positive in the sum, since $j \leq k$ and $i > k$. If we note that $\sigma_i' \geq \sigma_{i+1}'$ implies $\sigma_i'^2/(\sigma_j^2 - \sigma_i'^2) \geq \sigma_{i+1}'^2/(\sigma_j^2 - \sigma_{i+1}')^2$, then we conclude that the sequence is decreasing: $f_{k+1} \geq f_{k+2} \geq \cdots \geq f_n$. Finally we use $\sigma_1 \geq \sigma_j$ in all denominators in the sum (21) and obtain (20). □

If the difference between $\sigma_i'^2$ and σ_{n+1}^2 is very small compared to the value of σ_{n+1}^2, then f_i can become much larger than 1 for an $i \leq k$, as a consequence of (19). In particular, $f_k \gg 1$ if $\sigma_k'^2 - \sigma_{n+1}^2$ is very small compared to the value of σ_{n+1}^2.

This situation can happen in ill-posed problems if we let the truncation level k be too large. Due to the interlacing of singular values (4), the scenario when $f_i \gg 1$ (for some $i \leq k$) involves all of the following conditions:

- the smallest singular value of A is almost equal to the smallest singular value of $[A\ b]$; this means that $Ax \approx b$ is close-to-nongeneric (see the discussion on nongenericity after Algorithm 2, as well as (17, Th 3.11));
- the smallest singular value of A (and that of $[A\ b]$) is close to being multiple, with multiplicity r larger than $n - k$;
- i (with $i \leq k$) is also large enough such that it falls into the set of indices corresponding to the close-to-multiple singular values, *i.e.*, $n - r \leq i \leq k$.

Note that when there is elementwise additive i.i.d. random noise with large enough variance σ^2 present in A and b, the smallest singular value of A (and $[A\ b]$) becomes close-to-multiple, $\sigma_i' \approx \sigma_{i+1} \approx \sigma_{i+1}' \approx \cdots \approx \sigma_n' \approx \sigma_{n+1} \approx \zeta$, where the value ζ and the index i are determined by the variance of the noise and the singular value distribution of the noiseless data matrix. (See Figure 1.) Thus all the items above will hold true for some values of k and i.

On the other hand, if $\sigma_i'^2 - \sigma_{n+1}^2$ is not small compared to the value of σ_{n+1}^2 for any $i \leq k$, then we cannot use (19) as above to conclude that $f_k \gg 1$. We turn our attention to the bound (20); it shows that if V_{22}^k is close to zero, with $\|V_{22}^k\|$ smaller than σ_i'/σ_1 for some $i > k$, then $f_i > 1$. In particular, $f_{k+1} \gg 1$ if and only if $\|V_{22}^k\| \ll \sigma_{k+1}'/\sigma_1$. Thus a small $\|V_{22}^k\|$ (which may occur in close-to-nongeneric problems) implies a value of f_{k+1} greater than 1.

Example

Continuing with the noisy ill-posed example of §2.2, Figure 6 plots the TTLS filter factors, see (15), computed for a range of truncation levels k from 1 to n. Notice that for small truncation levels k ($k \leq 12$) the filter factors follow very well the characterization given by (4) (*i.e.*, the filter factors with indices $i = 1, \ldots, k$ stay close to 1 and the filter factors with indices $i = k+1, \ldots, n$ stay close to zero). However, when we gradually increase the truncation level k, there is a sudden point where this characterization no longer holds. From then on (*e.g.*, $k \geq 16$), the filter factors with indices i near the truncation level k increase dramatically, spoiling their regularizing property and illustrating that the ill-posedness comes into play.

3.4 Effective number of parameters, modified generalized cross validation and generalized information criteria

We have analyzed the properties of the filter factors for a fixed k; we have seen that for a k above a certain critical value the filter factors with indices nearby k are larger than 1. The idea is to identify the value of k where the filter factors change their nice behavior into the erratic growth behavior, since in this way we are likely to find a good truncation level for TTLS (and TSVD as well). Although there might be other ways to make use of the observed properties of the filter factors, we shall focus on the sum of the filter factors as a quantity of interest. For small values of k Lemma 1 implies that the sum will be almost equal to k; but for k above a certain critical value Lemma 2 shows that the sum will definitely be greater than k.

The sum of the filter factors is an important quantity for the general regularized solutions of the form (13). This sum corresponds to the *effective number of parameters* in the model, which is defined (*e.g.*, in (13)) as the trace of the *hat matrix*—the matrix that linearly transforms the noisy right-hand side into the regularized model b^{reg} (see (14)). The hat matrix is thus an $m \times m$ matrix H^{reg} satisfying $b^{\text{reg}} = H^{\text{reg}} b$. (For the model in (14) $H^{\text{reg}} = U' \text{diag}\{f_1, \ldots, f_n\} U'^{\top}$, and thus its trace is $\text{Tr}(H^{\text{reg}}) = \sum_{i=1}^{n} f_i$.)

The trace of the hat matrix appears in several model selection methods including the generalized cross validation (19, Chapter 4) and the generalized information criterion (12). The criteria (6) and (7) for TSVD, as well as the following criteria (22) and (23) for TTLS, are only special cases of these more general methods, where the trace of the hat matrix is replaced by simpler expressions that are suitable for the considered truncation methods, *i.e.*, by k or by the sum of the filter factors. (We refer also to (10, §7.4) for a discussion on GCV involving filter factors.)

In a linear regression model the number of effective parameters is exactly equal to the number of *free* model parameters; in TSVD, this number is k, the truncation level. Indeed, the TSVD solution $x_{\mathrm{TSVD},k}$, though of length n, corresponds to a corrected right-hand side $\widehat{b}_k := Ax_{\mathrm{TSVD},k}$, which lies in a k-dimensional linear subspace, and it is an orthogonal projection of b: $\widehat{b}_k = U'_k {U'_k}^\top b$. (Here, U'_k denotes the matrix of the first k left singular vectors of A.) The hat matrix in this case is simply $H_{\mathrm{TSVD},k} := U'_k {U'_k}^\top$, whose trace is equal to k. Thus, the number of effective parameters that describe the TSVD model is k.

We denote the effective number of parameters for the TTLS case, computed as the sum of the TTLS filter factors, see (15), by p_k^{eff}; this value can be plugged into the following modified criteria (compare with (6) and (7)):

- generalized cross validation:

$$\min_k \frac{\|Ax_{\mathrm{TTLS},k} - b\|^2}{\left(m - p_k^{\mathrm{eff}}\right)^2}, \tag{22}$$

- generalized information criterion

$$\min_k \ \log(\|Ax_{\mathrm{TTLS},k} - b\|) + \frac{p_k^{\mathrm{eff}}}{m}. \tag{23}$$

We have mentioned that the filter factors greater than 1 give a de-regularizing effect to the TTLS solution, in the sense that the SVD components corresponding to those filter factors are emphasized in the solution, instead of being damped. However, the fact that p_k^{eff} is greater than k when k reaches its critical value is beneficial for the two criteria above, since this induces a better defined minimum. Care should be taken in (22) that the values of k for which p_k^{eff} becomes greater than m are not taken into consideration. A practical choice (also used in our results in section 4) is to minimize (22) or (23) only for the integers k in the interval $[1, k_{\max}]$, where $k_{\max} \leq n$ is the maximum k such that $p_1^{\mathrm{eff}} \leq p_2^{\mathrm{eff}} \leq \cdots \leq p_{k_{\max}}^{\mathrm{eff}} \leq m$.

Note that the parameter p_k^{eff} can also be computed by using an alternative formula for the filter factors. When an $x_{\mathrm{TTLS},k}$ is already computed, its corresponding filter factors are just scaled linear coefficients in the linear expansion of $x_{\mathrm{TTLS},k}$ in the basis given by the right singular vectors of A. It is easy to see that from

$$x_{\mathrm{TTLS},k} = \sum_{i=1}^{n} f_i \frac{{u'_i}^\top b}{\sigma'_i} v'_i \tag{24}$$

we get

$$f_i = \sigma'_i \frac{{v'_i}^\top x_{\mathrm{TTLS},k}}{{u'_i}^\top b}. \tag{25}$$

4 Examples

We use simulated ill-posed problems such as the problem described in §2.2, from the Regularization Tools (9). However, not all ill-posed problems provided in the Regularization Tools can be solved satisfactorily using truncation methods. We explain this by the fact that not all types of ill-posedness are removable by simple truncation, and more complicated regularization methods should be used in those cases. The studied truncation methods favor solutions of small norms, which might not always be the most suitable property.

The results in this section test the performance of one of the truncation level selection methods, namely the generalized cross validation criterion proposed for TTLS in §3.4, (22). For the ill-posed example that we used from the beginning (see §2.2), the proposed GCV function (22) for choosing the truncation level is compared with the classical GCV criterion (6) that applies to the TSVD method. In the left plot of Figure 3, we notice that the modified GCV has a more clear minimum around the $k = 12$ value, which is actually the intuitively optimal value for this noisy problem—the k value that minimizes $\|x_{\mathrm{TTLS},k} - x_0\|$. (Recall that $x_{\mathrm{TTLS},k}$ is computed by (2) in Algorithm 2, and x_0 is the exact solution of the simulated example described in §2.2.) On the other hand, the classical GCV has a flat minimum; this would imply that choosing a larger k (*e.g.*, $k = 20$) is almost as good. But by checking Figure 5, it is obvious that $k = 20$ is a very bad choice.

The right plot of Figure 3 illustrates the comparison (in log scale) between the effective numbers of parameters that are computed in the two scenarios. For the classical GCV (6) the effective number of parameters for truncation level k is equal to k itself.

We have performed an extensive study to compare the GCV criterion proposed for TTLS and the classical GCV applied to the TSVD method. We used 8 problems from the Regularization Tools, namely:

1. `ilaplace(n,1)` 2. `ilaplace(n,3)` 3. `baart` 4. `shaw`

5. `phillips` 6. `foxgood` 7. `deriv2` 8. `wing`

We considered the case when both A and b contain additive noise or when only b contains additive noise; the relative noise level was 0.001. For both situations 100 noisy realizations were considered for each of the 8 example problems. Each problem was solved using TTLS and TSVD, and the truncation level was chosen with the corresponding GCV criterion, see (22) and (6) respectively. Moreover, we define for each simulation example the *best* truncation level as being the truncation level that provides the smallest error between the truncated and the true solution. Thus, for TSVD we obtain this best truncation level as the k that minimizes $\|x_{\mathrm{TSVD},k} - x_0\|$, while for TTLS we choose

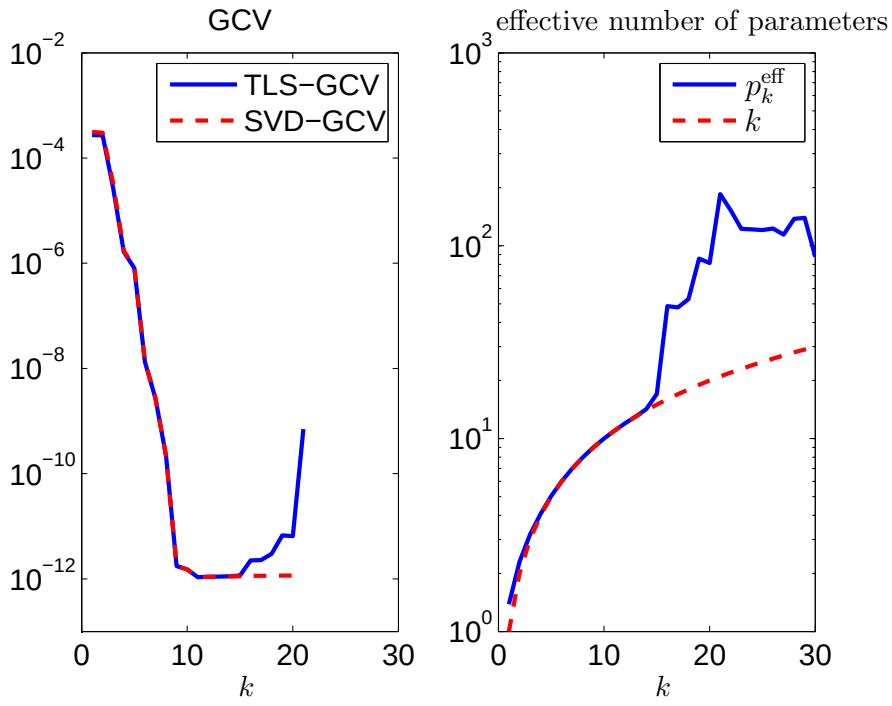


Fig. 3. The left plot shows the GCV function in the TTLS and the classical TSVD cases for the example in 2.2. The right plot compares the effective number of parameters in the TTLS case and in the TSVD case. Note that p_k^{eff} is approximately equal to k for a certain range of k values.

it by minimizing for k the value of $\|x_{\text{TTLS},k} - x_0\|$. In principle, the "best" truncation levels in the TSVD and TTLS case can be different, but they coincided in more than 90% of the simulations. The truncation levels chosen with the GCV criteria (in the TSVD and TTLS cases) are compared with the best truncation levels in Figure 4. The GCV-chosen truncation level coincides with the best truncation level, or it is at a distance no more than 2 to it, in the following percentages:

	TTLS	TSVD	TTLS	TSVD
noisy A and b	36%	26%	79%	65%
noise only on b	37%	32%	75%	70%
	coinciding		≤ 2 away	

In the neighborhood of the best k the GCV-chosen truncation levels are distributed equally above and below this value. However, there are some outliers that overestimate the best k level. Figure 4 illustrates this information by plotting the ratios between the truncation levels computed by GCV and the best truncation levels. The case with most outliers is when there is noise on A and b and we use TSVD.

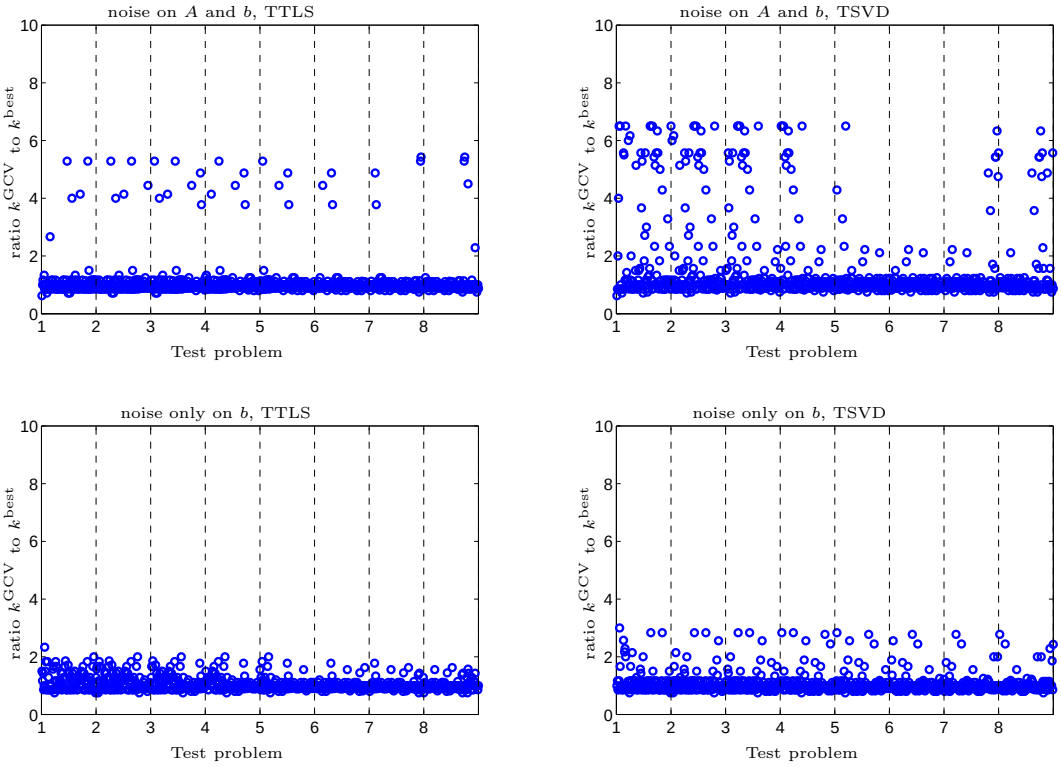


Fig. 4. Comparison between the two generalized cross validation (GCV) criteria, corresponding to the TTLS and the TSVD methods. For each test problem (from 1 to 8) we use 100 random noise realizations. Each plot shows the ratio between the value of the truncation level k^{GCV} chosen by GCV and the best k^{best} value minimizing the distance of the truncated solution to the true solution. Plotted values equal to 1 are thus optimal. The four plots correspond to: (top) noisy A and b, (bottom) noiseless A, noisy b; (left) $Ax \approx b$ solved with truncated total least squares; (right) $Ax \approx b$ solved with truncated singular value decomposition.

We mention briefly that the same simulations combined with Akaike's and the generalized information criteria (7) and, respectively, (23), instead of GCV, resulted in a very similar behavior. However, the case with noise on A and b solved by TSVD and AIC (corresponding to the upper right plot in Figure 4) showed even more outliers that overestimate the best truncation level.

5 Conclusion

We addressed approximate linear modeling problems in the case when the data comes from ill-posed problems, and we used truncation methods as a form of regularization.

We discussed model selection techniques for choosing truncation levels, and we emphasized that by considering the truncated total least squares problem we can exploit the following observation: when the truncation level is overestimated, the noise or the rounding errors take over the computed solution in a much higher degree than for the classical truncated SVD solution. The filter factors associated with the truncated total least squares solution are not necessarily equal to one and zero (as in the truncated SVD case). For this reason, they provide more information that can be used in criteria for choosing the truncation level; we illustrated the merits of a modified generalized cross validation method.

Although the modified generalized cross validation and generalized information criterion that we propose might be directly applicable for problem of moderate dimensions, where the full singular value decompositions of A and $[A\ b]$ can be computed, the next step is to embed these approaches in iterative methods appropriate for large-scale problems. The bidiagonalization algorithm for truncated total least squares presented in (4), or the reduction of data to a core problem proposed in (15), are good starting points for this goal.

Acknowledgments

The authors thank two anonymous referees for their valuable suggestions.

Dr. Diana M. Sima is a research assistant and Dr. Sabine Van Huffel is a full professor at the Katholieke Universiteit Leuven, Belgium.
This work was supported by: **Research Council KUL:** GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering, IDO 05/010 EEG-fMRI, several PhD/postdoc & fellow grants; **Flemish Government: FWO:** PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0360.05 (EEG, Epileptic), G.0519.06 (Noninvasive brain oxygenation), FWO-G.0321.06 (Tensors/Spectral Analysis), G.0341.07 (Data fusion), research communities (ICCoS, ANMMM); **IWT:** PhD Grants; **Belgian Federal Science Policy Office:** IUAP P6/04 ('Dynamical systems, control and optimization', 2007-2011); **EU:** BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST-2004-27214), FAST (FP6-MC-RTN-035801); **ESA:** Cardiovascular Control (Prodex-8 C90242)

References

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] R. Beran. Adaptive estimation of a mean matrix: Total least squares versus total shrinkage. Available from: http://www.stat.ucdavis.edu/~beran/tlsfin.pdf, 2006.

[3] P. Craven and G. Wahba. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, 31:377–403, 1979.

[4] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary. Regularization by truncated total least squares. *SIAM Journal on Scientific Computing*, 18(1):1223–1241, 1997.

[5] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223, 1979.

[6] G. H. Golub and C. F. Van Loan. An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17:883–893, 1980.

[7] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, third edition, 1996.

[8] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 32:561–580, 1992.

[9] P. C. Hansen. Regularization Tools, a Matlab package for analysis of discrete regularization problems. *Numerical Algorithms*, 6:1–35, 1994.

[10] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, 1998.

[11] P. C. Hansen, M. Kilmer, and R. H. Kjeldsen. Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT*, 46:41–59, 2006.

[12] S. Konishi and G. Kitagawa. Generalised information criteria in model selection. *Biometrika*, 83(4):875–890, 1996.

[13] J. E. Moody. The *effective* number of parameters: an analysis of generalization and regularization in nonlinear learning systems. In Moody, Hanson, and Lippmann, editors, *Advances in Neural Information Processing Systems*, volume 4, pages 847–854. Morgan Kaufmann, Palo Alto, 1992.

[14] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet. Math. Dokl.*, 7:414–417, 1966.

[15] C. C. Paige and Z. Strakoš. Core problems in linear algebraic systems. *SIAM Journal on Matrix Analysis and Applications*, 27(3):861–875, 2006.

[16] T. Reginska. A regularization parameter in discrete ill-posed problems. *SIAM J. Sci. Comput.*, 17(3):740–749, 1996.

[17] S. Van Huffel and J. Vandewalle. *The Total Least Squares Problem: Computational Aspects and Analysis*, volume 9 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1991.

[18] J. M. Varah. Pitfalls in the numerical solution of linear ill-posed problems. *SIAM Journal on Scientific and Statistical Computing*, 4(2):164–176, 1983.

[19] G. Wahba. *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics 59. SIAM, 1990.

[20] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Accoust. Speech Signal Processing*, ASSP-33(2):387–392, 1985.

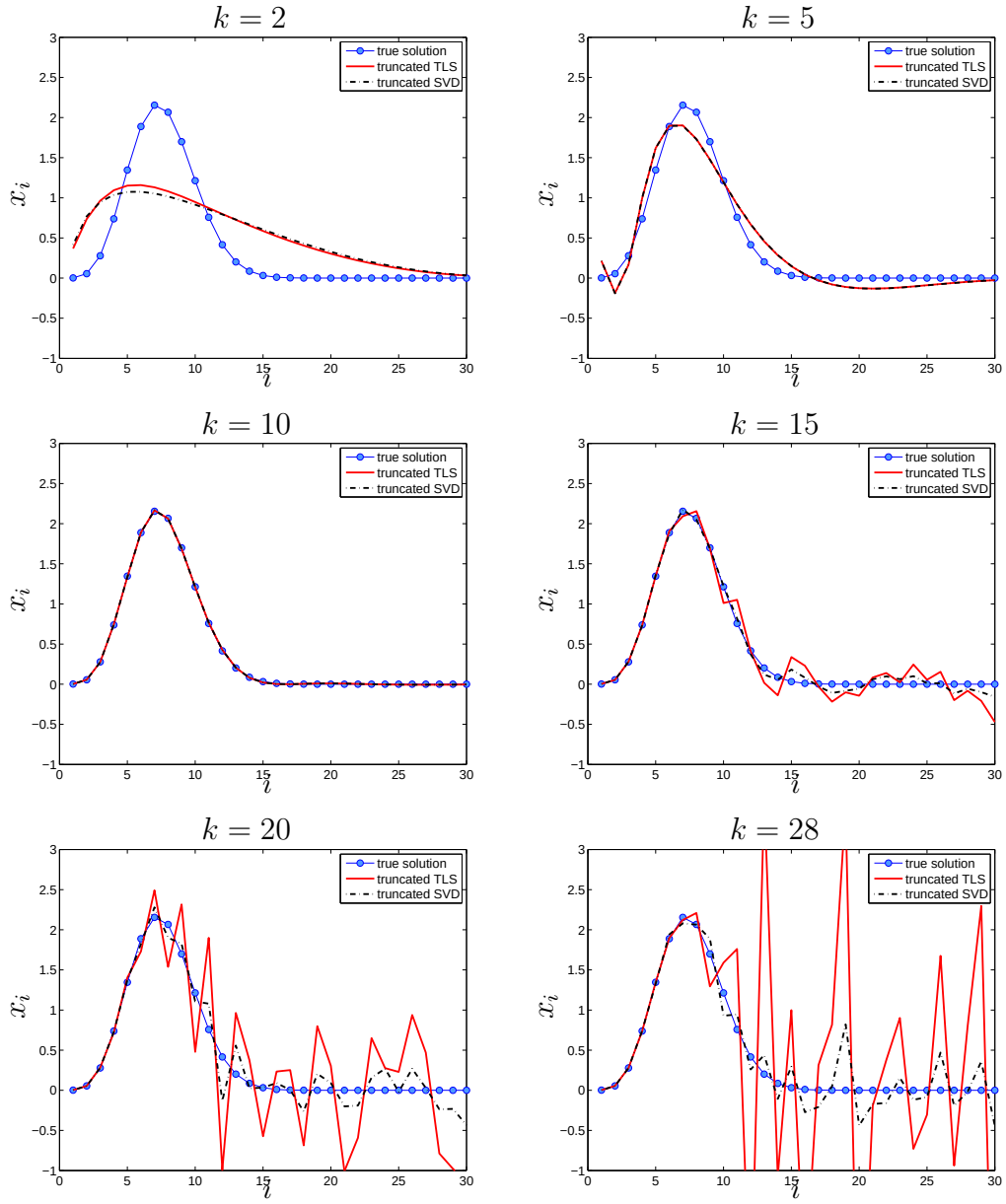


Fig. 5. Several truncated TLS and SVD solutions x_k, together with the true solution x_0, for the noisy simulation data described in §2.2. For both TTLS and TSVD solutions, the best truncation level is $k = 12$, minimizing $\|x_k - x_0\|/\|x_0\|$ to the value of 1e-3.

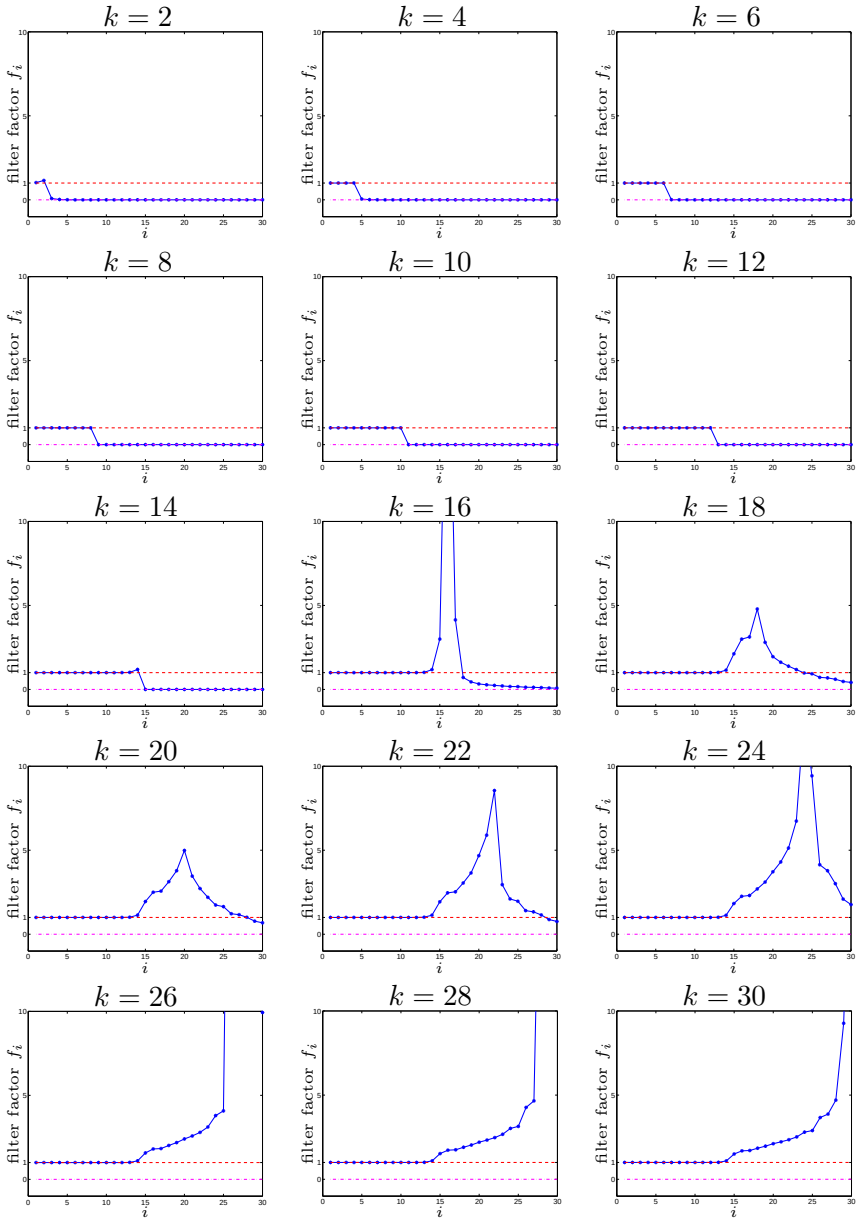


Fig. 6. The filter factors associated with the TTLS solutions for a representative range of k values between 1 and $n = 30$. The interrupted horizontal lines mark the zero and one levels. The numerical example is the same noisy example as in §2.2.